DOCUMENT RESUME

ED 387 699 CE 070 106

AUTHOR	Stoessiger, Rex; Pugh, Roy
TITLE	Professional Development for Program Evaluation: Evaluation for Adult Literacy and Basic Education Programs.
SPONS AGENCY	National Staff Development Committee for Vocational Education and Training, Chadstone (Australia).
REPORT NO	ISBN-0-642-22663-6
PUB DATE	95
NOTE	511p.
AVAILABLE FROM	Adult Basic Education Resource and Information Service, National Languages and Literacy Institute of Australia, Level 9, 300 Flinders St., Melbourne, Victoria 3000, Australia ($40 Australian).
PUB TYPE	Guides - Classroom Use - Teaching Guides (For Teacher) (052)
EDRS PRICE	MF02/PC21 Plus Postage.
DESCRIPTORS	*Adult Basic Education; Adult Educators; *Adult Literacy; Case Studies; Competency Based Teacher Education; Curriculum Guides; *Data Collection; *Evaluation Methods; Foreign Countries; Higher Education; Instructional Materials; Learning Activities; Learning Modules; Literacy Education; Professional Development; *Student Evaluation
IDENTIFIERS	Australia

ABSTRACT

This professional development program is designed to improve the understanding of Australian personnel working in adult literacy and basic education (ALBE) of the range of approaches and methods used to achieve different evaluation purposes. It is an adaptation of "A Guide to Program Evaluation" (included here), a detailed guide for people working in the ALBE field. The course is composed of six modules. The modules cover the following: introduction to evaluation and an overview of evaluation issues: evaluation approaches and methodologies; information collection alternatives; collecting and analyzing evaluative information; reporting and ensuring outcomes; and good practice in evaluation. Components of each module are as follows: an overview (duration; purpose; pre- or corequisites; summary of content; delivery; learning outcome; assessment criteria, conditions, and method; suggested resources); contents page; presenter's guidelines that provide a detailed description on how to present the units; transparency masters; and handouts. An appendix contains three case studies (evaluations) of ALBE projects that have been rewritten in an easily accessible style to provide examples of how evaluation projects have been conducted. (YLB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 387 699

CE070106

Professional Development for

PROGRAM EVALUATION

BEST COPY AVAILABLE

Evaluation for Adult Literacy and Basic Education Programs

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

● Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J Kindler

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A project commissioned by
National Staff Development Committee
for Vocational Education and Training

1995

Professional Development for Program Evaluation: Evaluation for Adult Literacy and Basic Education Programs is a project commissioned by:
the National Staff Development Committee
for Vocational Education and Training
Chadstone, Victoria

Published: 1995

Copyright: National Staff Development Committee
for Vocational Education and Training and the
Commonwealth of Australia

National Library of Australia ISBN: 0 642 22663 6

CONTENTS

INTRODUCTION 1

MODULE 1 Introduction to Evaluation and an Overview of Evaluation Issues 9

MODULE 2 Evaluation Approaches and Methodologies 51

MODULE 3 Information Collection Alternatives 93

MODULE 4 Collecting and Analysing Evaluative Information 117

MODULE 5 Reporting and Ensuring Outcomes 141

MODULE 6 Good Practice in Evaluation 165

APPENDIX Case Studies of Adult Literacy and Basic Education Evaluations 189

A Guide to Program Evaluation inside back cover

ACKNOWLEDGMENTS

NATIONAL PROJECT MANAGER

Liz Harris	National Staff Development Committee

PROJECT AUTHORS AND MANAGERS

Rex Stoessiger	Tasmanian Educational Consortium Ltd
Roy Pugh	Tasmanian Educational Consortium Ltd

OTHER PERSONNEL

A wide cross section of personnel gave input to this project. They included:

Steering Committee Members

Sue Shore	South Australia
Jan Kindler	Victoria
Susan Garside	Queensland
Liz Parkinson	NSW
Nancy Veale	ACT
Jim Thompson	Western Australia
Lee Clarke	Northern Territory

ALBE Practitioners who participated in the trial presentations and made many helpful suggestions about the course

Olive Briscoe, Jan Kindler and Liz Harris who made detailed contributions to the materials

PRODUCTION

Word Processing and Layout
Tasmanian Educational Consortium Ltd, Hobart

Editing and Production
Jeannette Thiering, Positive Education Services, NSW

Cover Design
Intelink, Woolloomooloo, NSW

INTRODUCTION

In 1993 the Commonwealth Department of Education Employment and Training commissioned the production of a detailed guide to evaluation for people working in the Adult Literacy and Basic Education (ALBE) field. The result was the production of a very comprehensive text, *A Guide to Program Evaluation*, written by Faye Lambert and John Owen in conjunction with Sharon Coates and Julie McQueen (DEET, 1994).

Subsequently the National Staff Development Committee (NSDC) commissioned Rex Stoessiger and Roy Pugh of the Tasmanian Educational Consortium Ltd to adapt the Guide to become a professional development course. As a result this course, *Professional Development for Program Evaluation*, was written. It was trialled with ALBE personnel in a variety of Australian states and sections were then redrafted.

1 COURSE AIMS

This professional development program aims to:

- improve participants' understanding of the range of approaches and methods used to achieve different evaluation purposes
- provide practical training to assist participants, at any level within ALBE operations, to undertake evaluations that will lead to improved efficiency and effectiveness
- point to ways of making evaluation techniques an integral part of program administration and implementation
- demonstrate the application of evaluation processes within the ALBE field and, in particular, focus on the selection and use of performance indicators, reflecting the operational functions at national, state or territory, regional and provider level.

(Adapted from *A Guide to Program Evaluation*, pp. 1–2)

2 COURSE STRUCTURE

The course is composed of six modules. Module 1 is a core module.

Module 1	Introduction to Evaluation and an Overview of Evaluation Issues
Module 2	Evaluation Approaches and Methodologies
Module 3	Information Collection Alternatives
Module 4	Collecting and Analysing Evaluative Information
Module 5	Reporting and Ensuring Outcomes
Module 6	Good Practice in Evaluation

See also Item 5 of this Introduction which outlines different ways of structuring the course to meet the needs of particular participants.

In addition to the six modules, participants who are doing the full course will undertake an evaluation project at their own ALBE workplace, called the Participant's Evaluation Project.

PARTICIPANT'S EVALUATION PROJECT

The overall course is designed so that participants can undertake an evaluation project as an action learning task, between sessions. This allows participants to use the ideas developed during the course in practical situations while they are fresh in their minds.

At the start of each new session participants have the opportunity to report on the progress of their project to date, allowing them to get assistance, and providing useful insights for other course members on the practical consequences of conducting program evaluations.

Participants are encouraged to attempt a *small* evaluation project which contributes to their normal work. The 'workplace' for the project may be a teaching/learning program in any ALBE context. This could, of course, include an ALBE program conducted at an industrial site or in a commercial enterprise.

TOTAL NOMINAL TIME

Each of the six modules takes about 3 hours of workshop time.
The Participant's Evaluation Project will take about 15 hours extra.

Total: 33 hours for the full course

3 ENTRY REQUIREMENTS

This program is for all personnel working in ALBE and related areas.

- It is for people operating at national, state or territory, regional and local levels who need to undertake evaluations as part of their work.
- It is for use by systems, departments, organisations and work groups.

The course would be useful in a wide variety of areas to people who are involved with evaluation in general, and program evaluation in particular. The examples given largely relate to ALBE, but could easily be supplemented or replaced by examples from other areas as the basic approach applies to evaluation in any field.

No background knowledge of evaluation is required to start the course, but experience with evaluations will assist participants.

7

4 RECOGNITION OF PRIOR LEARNING

Participants who wish to have prior learning recognised instead of completing this program may apply to the course coordinator of each module for assessment of current competence against the learning outcomes identified in the program.

5 DELIVERY

MODES OF DELIVERY

The preferred approach in this professional development program is a highly interactive one. Adult learners have much to contribute to their own professional development and training. Good courses both respect and utilise the thinking of participants.

As a result the course is a balance of relatively small amounts of presenter inputs with longer periods spent in organised group discussions. The use of a wide variety of group sizes is suggested, including pairs and small groups; and the whole group for brainstorming and discussions.

Often participants report one of the major benefits of attending professional development sessions has been spending time with colleagues. Accordingly strategies such as group investigations with all participants reporting to new groups (see Additional Note 1.1 'Modified Jigsaw Movement') have been incorporated to ensure that participants have opportunities to work in a variety of groups as part of the course.

RESTRUCTURING THE COURSE

In Unit 1.4 participants undertake a self-evaluation of their prior learning about evaluation. The self-evaluation may suggest the need to re-structure or re-focus aspects of the course. Presenters will, ideally, be flexible enough to do this. For example, if all participants reveal a thorough knowledge of evaluation fundamentals then less time might be spent on Module 1 and more time devoted to subsequent modules.

NUMBER OF PARTICIPANTS

To enable effective small and whole group interaction, there should not be more than 20 participants.

PRESENTERS' QUALIFICATIONS

Presenters should:

- be experienced in delivering professional development programs
- have an excellent knowledge of the theory and application of evaluation
- be up-to-date with the broad issues affecting the ALBE field.

PROGRAM FLEXIBILITY

Full Course

The six modules are designed to fit together to provide a comprehensive course in evaluation. Each module takes about half a day so the whole course could run as six half day sessions or as a three day course. In either of these formats participants can be asked to conduct their evaluation project as part of the course and report on its progress during subsequent sessions.

Short Courses

While designed to fit together, units can also be used independently. For example, Module 1 can stand alone as an introduction to evaluation. Similarly those with a reasonable understanding of evaluation issues who wish to find out more about a range of data collection alternatives could use Module 3 only.

Modules 1 and 2 fit together well, as do Modules 3 and 4, and also Modules 5 and 6. Any unit prerequisites are mentioned in the module overviews.

Short courses, without time spent back at work, lose the option of undertaking the Participant's Evaluation Project as part of the course.

Composite Courses

Composite courses can be constructed by making careful selections from several modules. For example, a one day course for people with some background knowledge of evaluation might consist of the units shown in the following table.

Unit	Title
1.8	Planning an evaluation
1.9	The evaluation question
1.10	Evaluation for different purposes
2.6	Evaluation for program improvement
2.7	Evaluation for program justification and accountability
3.4	Information collection alternatives
4.4	Sampling
5.5	Ensuring the impact of evaluation findings

Such combinations need to be made with care to ensure some connection between units. With this in mind it is also possible to negotiate with participants to select units which meet their specific needs.

6 REQUIREMENTS FOR THE AWARD

Learning outcomes, assessment criteria, conditions and assessment methods are listed for each module.

Those who attend the full program and complete the assessment task to the required standard will receive a Statement of Attainment, provided that the course has been accredited in their state or territory.

Otherwise the award will be a Statement of Achievement outlining the sections of the course completed.

7 RECOGNITION OF THE AWARD

There have been no formal arrangements for recognition of the award.

Participants may choose to present the Statement of Attainment, or Statement of Achievement, for advanced standing towards courses offered by tertiary and post-secondary institutions.

8 RELATIONSHIP TO COMPETENCY STANDARDS

The course is designed to correlate with competency statements developed by the Department of Education Employment and Training's International Literacy Year project, 'What Is a Competent Adult Basic Education Teacher?' In 1990 DEET set up this ILY project which was conducted by the University of Technology, Sydney. An interim report was published, *What is a Competent ABE teacher?* and in 1993, a final report, *The ABE Profession and Competence: Promoting Best Practice*. The UTS work was developed with cross sector input from all states and territories and provides the only framework available for ALBE personnel.

The competency statements relevant to this course are found in chapter 4, section 4.5, of *The ABE Profession and Competence: Promoting Best Practice*, p. 28.

Unit 5: Evaluation of Programs

Element 5.1
Examines existing models of evaluation to adopt and incorporate appropriate evaluation strategies, and uses the evaluation of courses for revision, planning and development.

Element 5.2
Reports to relevant authorities on program effectiveness.

9 RESOURCES

THE MATERIALS

Five types of materials are included in this package to assist with the delivery of this course.

1 **Presenter's Guidelines** which provide a detailed description on how to present the course units. Presenter's Guidelines are identified by the letters **PG** highlighted in the header of each page.

2 **Overhead Projector Transparencies** which are identified by the highlighted letters **OHP**.

3 **Handouts** for participants, with the letter **H** highlighted.

4 **Additional Notes** which provide advice to the presenter about aspects of the course and its presentation and have the letter **AN** highlighted.

5 **Case Studies** which are placed in an appendix.
Three published evaluations of ALBE projects have been re-written in an easily accessible style to provide case studies of how evaluation projects have been conducted. The case studies are referred to at key points in the Presenters' Guidelines, but can also be copied and distributed to participants as supplementary readings.

A Guide to Program Evaluation

The above materials, except for the case studies, are adapted from *A Guide to Program Evaluation*, referred to throughout this document as 'the Guide'.

One copy of *A Guide to Program Evaluation* is supplied with each of these packages.

The Presenter's Guidelines refer to the pages of the Guide that the presenter will need to copy. It is essential that presenters study the Guide thoroughly and use it alongside this professional development course. The Guide is much more comprehensive than this document, as only some parts of the Guide could be condensed into a course of reasonable length. Hence the Guide is also recommended for further reading, along with additional references for participants who wish to continue learning about evaluation.

The Commonwealth of Australia copyright specifications for *A Guide to Program Evaluation* allow it to be 'reproduced in whole or in part for study or training purposes subject to inclusion of an acknowledgment of the source; it is not to be copied for commercial use or sale' (*The Guide*, p. i). Presenters must respect these conditions.

10 ASSESSMENT

There are two assessment methods for participants doing the full course.

1 They will be assessed according to the level of their participation in the workshop activities.

2 They will conduct a Participant's Evaluation Project which includes:
  - developing an evaluation plan for a program at their own workplace
  - collecting the information required by the evaluation plan
  - processing the information collected
  - reporting on the evaluation findings in some suitable format.

In Module 6 an interim report will be given to the group by each participant doing an evaluation project. Within one month of the program a final project report should be sent to the coordinator who will be responsible for assessing whether participants have satisfactorily completed their evaluation project.

The coordinator and/or presenter will assess whether participants have achieved the learning outcomes for this course.

The coordinator and/or presenter will ensure that a Statement of Attainment (or Statement of Achievement, if appropriate) is sent to each successful participant and will also handle requests for recognition of prior learning.

MODULE 1

INTRODUCTION TO EVALUATION AND OVERVIEW OF EVALUATION ISSUES

14

13

OVERVIEW OF MODULE 1

1 AN INTRODUCTION TO EVALUATION AND AN OVERVIEW OF EVALUATION ISSUES

2 NOMINAL DURATION

Three hours

3 PURPOSE

This module will provide an overview of the major issues associated with conducting an evaluation.
It introduces some important evaluation ideas and terms.
It prepares participants to begin planning an evaluation for themselves.

4 RELATIONSHIP TO COMPETENCY STANDARDS

See Item 8 of the Introduction, Relationship to Competency Standards.

5 PRE- OR CO-REQUISITES

None

6 SUMMARY OF CONTENT

An introduction to evaluation and an overview of the major issues associated with conducting an evaluation.
An outline of the procedures for planning an evaluation.
Planning an evaluation of a program at the participant's own workplace.

7 DELIVERY

Interactive workshop.
Small and large group discussions.
Group investigations with all participants reporting their findings to new groups.

8 LEARNING OUTCOME

On completing this module participants will be able to start to plan an evaluation for a particular purpose in a specific context (e.g. the Assessment Task for the course).

9 ASSESSMENT CRITERIA

On completing this module participants will be able to:

- define evaluation and list the major components of an evaluation
- give reasons why the context of an evaluation is important
- list the questions to be answered when planning an evaluation

- explain why clarifying the purposes of an evaluation is important
- begin to plan an evaluation.

10 ASSESSMENT CONDITIONS

The coordinator and/or presenter will support the learner in the workshop activities and in the development of the Participant's Evaluation Project and will provide access to all necessary resources.

There will also be cooperative teaching/learning between participants as peers.

11 ASSESSMENT METHOD

Participants are assessed in two ways.

- They must show through discussion and other workshop activities that they have a sound concept of evaluation.
- Their Participant's Evaluation Project, when completed, must show evidence that they can plan an evaluation of an actual ALBE program .

12 SUGGESTED RESOURCES

Lambert, F. & Owen, J. 1994, *A Guide to Program Evaluation*, DEET

MODULE 1 CONTENTS

1 Welcome and introductions

2 Overview of the course

3 Agenda for the first workshop

4 Self-evaluation of what participants already know about evaluation

5 Definition of evaluation

6 Components of an evaluation

7 Importance of context

8 Planning an evaluation

9 The evaluation question

10 Key Aspect 1: Evaluation for different purposes

11 Using the ideas from the module to plan the Participants' Evaluation Projects

12 Session evaluation and close

PG OHP H AN

CONTENT	PRESENTER'S GUIDELINES	RESOURCES
SESSION 1 (about 3 hours)		
1.1 Welcome and introductions (10 minutes)	Welcome participants and undertake an introductory activity which provides an opportunity for all participants to speak to the WHOLE GROUP. *For example, participants introduce themselves, give a brief indication of their background and why they are attending the course.*	
1.2 Overview of the course (5 minutes)	Distribute Handout 1.1 which lists the purposes of the course (adapted from the Guide, pp. 1–2).	Handout 1.1 'The purpose of the course'
	Provide an overview of the course (Transparency 1.1) and encourage any initial questions.	Transparency 1.1 'Course overview'
1.3 Agenda for the first workshop (5 minutes)	Outline the purpose of Module 1 and the agenda for this workshop using Transparency 1.2.	Transparency 1.2 'Agenda —Module 1'
	Mention the Participants' Evaluation Projects briefly at this stage if they are to be conducted as part of this delivery of the course.	

1.4 **Self-evaluation of what participants already know about evaluation** (20 minutes)	Introduce the idea of self-evaluation and allow brief discussion to clarify what is meant.	
	Ask the group to work in PAIRS and to list all the things they already know about evaluation ready for sharing with the whole group.	blank OHTs or butchers paper
	Arrange WHOLE GROUP sharing of prior knowledge about evaluation.	***Note**: You may need to restructure the course at this stage to meet the particular needs of participants.*
	List key points and issues for further examination on white board/butchers paper etc.	
1.5 **Definition of evaluation** (15 minutes)	In the WHOLE GROUP agree on some of the requirements for a working definition of evaluation, e.g. 'involves the collection and analysis of information', 'making judgments or decisions', 'possibly based on criteria'.	
	Display Transparency 1.3 as a working definition of evaluation for this particular course.	Transparency 1.3 'A definition of evaluation'
	Draw participants attention to the list of terms defined in Handout 1.2.	Handout 1.2 'Definitions of evaluation terms'
	Encourage the group to ask questions and explore the ideas raised.	

PG OHP H AN

1.6 **Components of an evaluation** (10 minutes)	Display the stages of an evaluation as shown in Transparency 1.4: • Planning • Data Collection • Data Processing • Making Judgments • Reporting Outline the stages. Point out that the middle three stages are sometimes collapsed and called 'information handling' or something similar. Encourage participants to ask questions to help them understand the components of an evaluation.	Transparency 1.4 'The stages of an evaluation' *The components of an evaluation are discussed in Chapter 2 of the Guide, 'The Three Stages of Evaluation'. (The Guide, pp. 5-7)* *Making judgments is discussed in more detail in Unit 3.3 'Key Aspect 3: Making judgments'*

27

23

1.7 Importance of context (20 minutes) Allow about: – 5 minutes in the initial groups and – 10 minutes in the reporting groups.	Introduce the idea that an evaluation always occurs in a broad context: – organisational – economic – social – administrative and/or political. Explain that this will affect how the evaluation is conducted. Mention the importance of identifying the stakeholders in an evaluation. ◆ **The Evaluation Case Studies** in the appendix describe the contexts and stakeholders of each evaluation. GROUP INVESTIGATION–JIGSAW MOVEMENT	*Consult Additional Note 1.1 'Modified jigsaw movement' for a more complete explanation of the grouping and regrouping used in this section.*
	Place participants in FOUR EXPERT GROUPS. Assign each group the task of investigating how ***one*** of the four contextual influences (broad organisational, social, economic, administrative and/or political) might influence a typical ALBE evaluation. (See Additional Note 1.2)	Transparency 1.5 'Modified jigsaw movement'
	Re-group into REPORTING GROUPS containing one member from each expert group. (Display Transparency 1.5 to show the movement.) Participants now report on their findings to the new groups.	*Use Additional Note 1.2 'An example of the importance of context'*
	When participants come together as a WHOLE GROUP, encourage them to discuss any remaining issues briefly.	

PG OHP H AN

1.8 **Planning an evaluation** (5 minutes)	Introduce planning an evaluation as the first stage of an evaluation. (Refer back to Transparency 1.4 .)	
	Display the simplified approach to planning an evaluation in Transparency 1.6.	Transparency 1.6 'Planning an evaluation'
	Outline how the different questions drive the planning process. For example, it is essential to identify the evaluation question or questions (what is being evaluated?) at an early stage and ensure that this is generally agreed by all those involved.	
	Stress that • before you decide on an evaluation methodology you need the answers to the first four questions: What? Why? Who's involved? Who's it for? and • the methodology should come from the answers to these four questions.	**The first 4 questions:** **What**…? **Why**…? **Who**'s involved? **Who**'s it for?
	Encourage some discussion on planning evaluations using the recommended approach. The discussion might include: • the importance of determining what is to be evaluated and why • the use of the term 'stakeholders' to describe who has ownership of the program being evaluated (see the Guide, p. 26 for a definition of stakeholders) • why people often jump to the methodology before determining the preceding parts of planning an evaluation.	◆ **The Case Studies** give examples of evaluations described under similar headings.

1.9 **The evaluation question** (10 minutes)	Stress the following points: • it is essential to clarify and secure agreement about the evaluation questions at an early stage, as mentioned above (in Unit 1.8) • sometimes this will take time • program managers will typically have a rough idea about what needs evaluating, but will need an opportunity to clarify their ideas • the evaluator will need to help the manager or team focus on what really is important because – the evaluation questions will often be presented in a very broad and general way, or – an impossibly large list will be suggested for evaluation.	
	Ask participants, as INDIVIDUALS, to make some notes on the following question (Transparency 1.7): 'An ALBE project management team wants to include "everything" in the evaluation but resources are limited. How would you go about encouraging them to focus on particular evaluation questions?'	Transparency 1.7 'The evaluation question'
	Form PAIRS and ask individuals to share their notes with a partner.	
	Some suggestions are contained in Additional Note 1.3.	*Additional Note 1.3* *'Suggestions for focusing on particular evaluation questions'*
	* Examples of evaluation questions are included in the **Case Studies**.	

PG OHP H AN

1.10 **Key Aspect 1: Evaluation for different purposes** (25 minutes)	Explain to the WHOLE GROUP that evaluation can be conducted for many different purposes. Ask the group to brainstorm all possible reasons for an evaluation.	*Five basic purposes for evaluation are identified in the Guide, p. 5.*
	List all reasons offered on the whiteboard, including the more cynical ones. You may add some reasons of your own.	• Examples of evaluations for different purposes are included in the **Case Studies**.
	Group the reasons into categories and use them to draw attention to the principal reasons for evaluating. Handout 1.3 has four principal purposes as column headings and it may be possible to identify these purposes from the brainstorm. However there is no need to force this as there are many and varied reasons why people conduct evaluations.	Handout 1.3 'An overview of evaluation purposes' ***Note 1****: If possible enlarge this to A3 size.*
	Ask participants in PAIRS to examine Handout 1.3, 'An overview of evaluation purposes'.	
	Ask pairs to examine: • the different questions typically asked when evaluating for different purposes (the TYPICAL QUESTIONS section of the handout). and • the focus questions on Transparency 1.8.	***Note 2:*** *Handout 1.3 is quite complex and participants will need time to study it.* Transparency 1.8 'Focus questions'
	Allow pairs to share their results briefly with the whole group.	

36

1.11 Using the ideas from the module to plan the participants' workplace evaluation projects (10 minutes)	Point out that if a workplace project is to be undertaken as part of the course, the first thing for participants to do, back at work, is to plan an evaluation of a program that is important to them. Make it clear that by next session they should have documented: – the evaluation question – why they are conducting the evaluation – who is involved – the context of the evaluation, including stakeholders. *Participants should also negotiate any agreements required to be made in their own workplace to allow the conduct of the evaluation.* *This could include obtaining a supervisor's permission or agreement from colleagues to have their work evaluated.*	***Note:** This activity could also be undertaken as an exercise during the session. If so participants could work in groups and plan the evaluation of a program of mutual interest, such as an ALBE course. This would be particularly important for participants undertaking a short course which does not include the workplace project.*
1.12 Session evaluation and close (15 minutes)	Discuss with the WHOLE GROUP the likely purposes for conducting a course evaluation at this stage. For example, – to help improve future sessions – to monitor implementation – to gauge client satisfaction.	***Note:** If two modules are being presented in the one day this section can be deferred to the end of the day.*
	Ask participants to evaluate this workshop for future improvement by filling in Evaluation Form 1 and returning it before leaving. *Each module has a different type of evaluation form. Participants could appraise them critically to help them learn about evaluation methods.*	Handout 1.4 'Evaluation Form 1'

COURSE OVERVIEW

1 Introduction

2 Approaches and methodologies

3 Information collection

4 Collecting and analysing information

5 Reporting and ensuring outcomes

6 Good practice in evaluation

Transparency 1.1

AGENDA – MODULE 1

1. Welcome and introductions
2. Overview of the course
3. Agenda for the first module
4. Recognition of prior learning
5. A definition of evaluation
6. The components of an evaluation
7. The importance of context
8. Planning an evaluation
9. The evaluation question
10. Key Aspect 1:
    Evaluation for different purposes
11. Planning an evaluation project
12. Session evaluation and close

Transparency 1.2

A DEFINITION OF EVALUATION

Evaluation is:

the collection and

the analysis of information

in order to

facilitate

informed decision making

from *The Guide*, p. 1

Transparency 1.3

THE STAGES OF AN EVALUATION

1 Planning

2 Data collection

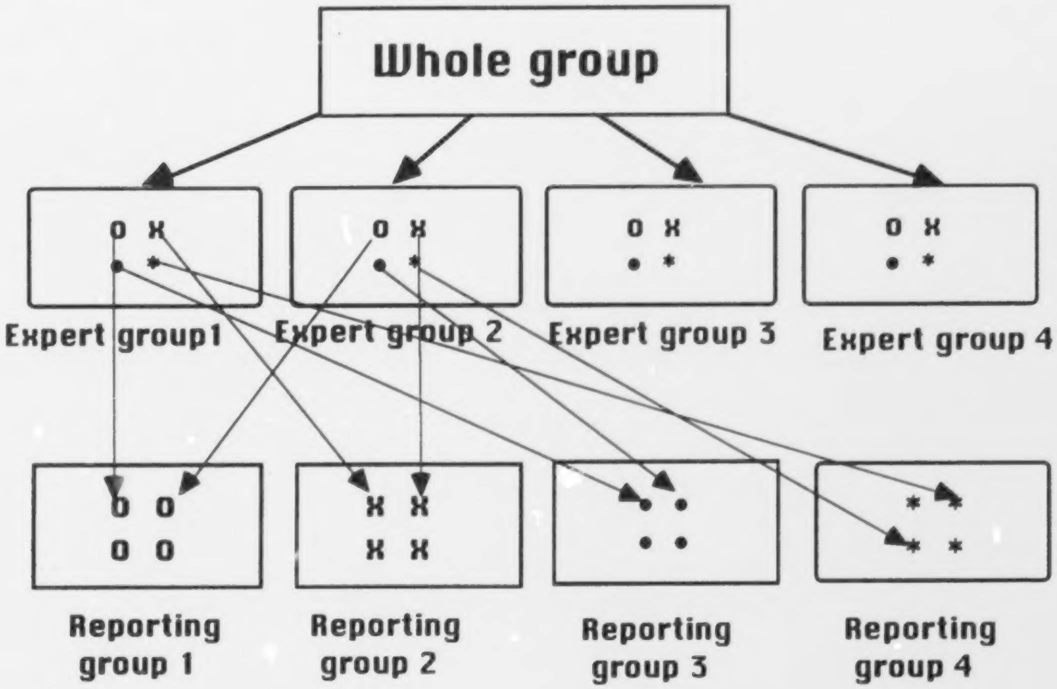
3 Data processing

4 Making judgments

5 Reporting

Transparency 1.4

MODIFIED JIGSAW MOVEMENT


Whole group
o x
• *
Expert group1
Expert group 2
Expert group 3
Expert group 4
o o
o o
x x
x x
• •
• •
* *
* *
Reporting group 1
Reporting group 2
Reporting group 3
Reporting group 4

Adapted from a diagram in Bennett, B., Rolheiser-Bennett, C. & Stevahn, L. 1991,
Cooperative Learning: Where heart meets mind,
Educational Connections, Toronto, Ontario, p. 219

Transparency 1.5

PLANNING AN EVALUATION

What is to be evaluated?

Why?

Who is involved?

Who is the evaluation for?

How will the program be evaluated?

When?

Who by?

Transparency 1.6

THE EVALUATION QUESTION

An ALBE project management team wants to include ‘everything’ in the evaluation, but resources are limited.

How would you go about encouraging them to focus on particular evaluation questions?

Transparency 1.7

FOCUS QUESTIONS

In what ways are different groups (e.g. funding agencies, managers and practitioners) interested in different purposes for an evaluation?

What might be the consequences of these differences?

Transparency 1.8

THE PURPOSE OF THE COURSE

The course is designed to:

- **improve** your understanding of the range of approaches and methods used to achieve different evaluation purposes
- **provide** practical advice that will assist you, at any level within ALBE operations, to undertake evaluations that will lead to improved efficiency and effectiveness
- **point to ways** of making evaluation techniques an integral part of program administration and implementation
- **demonstrate** the application of evaluation processes within the field of adult literacy and basic education, and in particular, focus on the selection and use of performance indicators, reflecting the operational functions at national, state/territory, regional and provider level.

Adapted from *The Guide*, pp. 1–2

Handout 1.1

DEFINITIONS OF EVALUATION TERMS

Appraisal is evaluating an individual's performance or work.

Assessment is collecting information in order to make judgments about learning, competency or attainment.

Effectiveness is the extent to which a program is satisfying the purpose for which it was established.

Efficiency is the economic use of inputs to achieve the program's outputs.

Evaluation is the collection and the analysis of information in order to facilitate informed decision making.

Performance indicators are simple statistics designed to give summary information which allows managers and others to make judgments about the effectiveness of programs. They are usually expressed as an index, rate or other numerical figure.

Performance information is any information (both quantitative and qualitative) which allows judgments to be made about the achievements of a program.

Reliability of an approach is its ability to obtain the same result on repeated occasions.

Self-evaluation is the evaluation of a program by the people responsible for planning and implementing it.

Stakeholders are people or groups who are significantly affected by the actions of a program or who impact on the program in important ways.

Validity of an approach is its ability to measure what it is intended to measure.

Handout 1.2

AN OVERVIEW OF EVALUATION PURPOSES

Handout 1.3

PURPOSES OF EVALUATION	**PROGRAM DEVELOPMENT** to make decisions about how best to develop a program to meet the needs, eg training needs analysis	**PROGRAM CLARIFICATION** to clarify the logic of a program to assist in the effective delivery of outcomes	**PROGRAM IMPROVEMENT** to improve the implementation of a program	**PROGRAM JUSTIFICATION / ACCOUNTABILITY** to monitor program delivery and outcomes, to determine program worth	
				MONITORING EVALUATION (accountability)	IMPACT EVALUATION (justification of worth)
FOCUS	context	design	delivery	outcomes / delivery	outcomes / delivery
TIMING	before	during	during	during	after
STAGE OF PROGRAM DEVELOPMENT	none	development	development	established	established
TYPICAL QUESTIONS	• What is the ideal or desired state of affairs? • What is the actual state of affairs? • Are there discrepancies between the desired and the actual? • What are the reasons for the discrepancies or needs?	• What are the intended outcomes of the program and how is the program designed to achieve them? • What program elements need to be changed to maximise its potential to achieve the intended outcomes? • What aspects of the program should be considered in a program evaluation? • Which of these aspects are amenable to evaluation? • What aspects require further attention before being evaluated at a later date?	• How is this service or activity going? • Is it working? • How is it affecting the target group or clients? • What specific aspects need improvement?	• Is the program reaching the target population? • Is it being implemented in ways specified? • Is it effective? • How much does it cost? • What is the cost relative to effectiveness?	• What are the 'real goals' of the program and to what extent are they attained? • What are the intended and unintended outcomes of the program? • How do differences in implementing the program affect the outcomes? • Does the program meet identifiable needs?
TYPICAL DECISIONS	• Should the development of a program go ahead? • What needs should be given priority for action through an intervention or program?	• Is there a need to update or revise the plan or logic of the program currently being delivered?	• Should there be changes in the way the program is delivered? • Should policy associated with the program be reviewed or updated? • Should the program be modified so that it is delivered more efficiently?	• Is the funding allocated to the program justified in terms of what criteria? • Can the public have confidence in the services offered through the program? • Can praise or blame for the success or failure of the program be allocated to program personnel? • Should the program be discontinued because of its failure to provide adequate levels of outcomes?	
ALBE EXAMPLES	*The Guide* pp. 19-20, pp. 22-23	*The Guide* pp. 29-30, pp. 33-34	*The Guide* pp. 38-40, pp. 42-43	*The Guide* p. 57, pp. 59-60, pp. 63-90	

EVALUATION FORM 1

Please fill in the evaluation form to help improve subsequent sessions.

1 What helped your learning in the session? ____________________________

____

____

2 What suggestions do you have for improving subsequent sessions?

____

____

____

____

Please rate the following aspects of the session on a 1 to 4 scale
with **1** meaning **very good** and **4** meaning **very poor**.

	1	**2**	**3**	**4**
3 The session content	☐	☐	☐	☐
4 The style of presentation	☐	☐	☐	☐
5 Depth of material covered	☐	☐	☐	☐
6 The type of activities	☐	☐	☐	☐
7 The support materials	☐	☐	☐	☐
8 Interaction with other participants	☐	☐	☐	☐
9 Getting what you came for	☐	☐	☐	☐

10 Any other comments? ____

____

____

____

____

____

Handout 1.4

ADDITIONAL NOTE 1.1
MODIFIED JIGSAW MOVEMENT

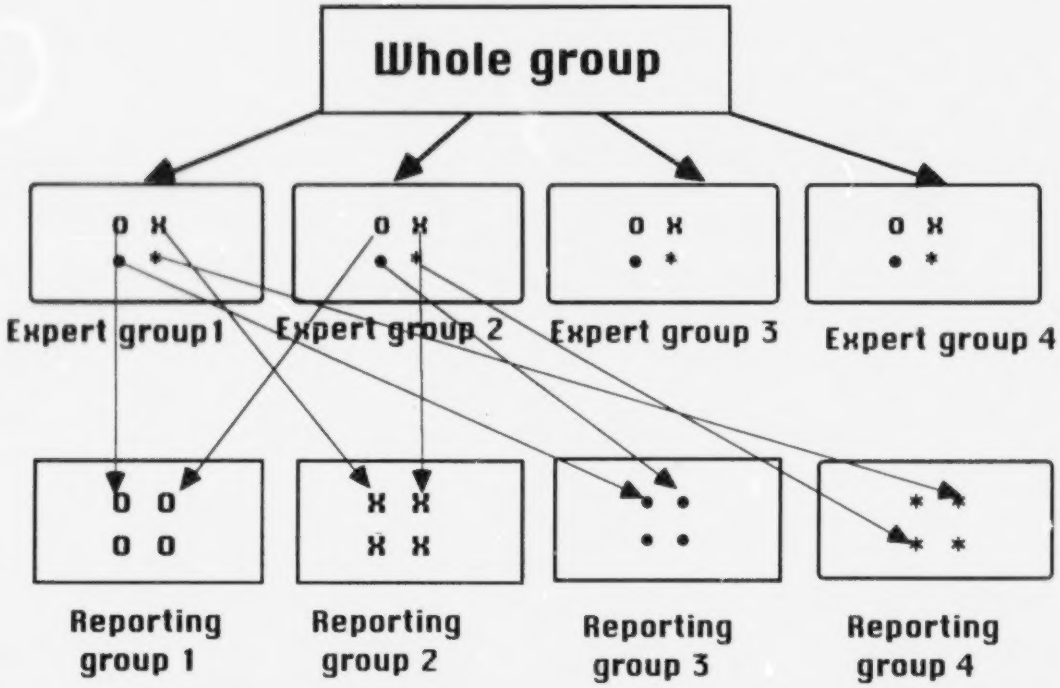
A common strategy used in this course is a group investigation in which participants are grouped in different ways in a movement based on the Jigsaw structure used in the cooperative learning literature (Bennett, B., Rolheiser-Bennett, C. & Stevahn, L. 1991, *Cooperative Learning: When Heart Meets Mind,* Educational Connections, Toronto, p. 219).

The structure used here is a two step process.

1 Divide participants into groups of roughly equal numbers. These are the expert groups. In these groups participants investigate a particular topic and prepare to report, individually, to new groups who have not studied this particular topic.

2 Move expert group members to new groups for reporting on their findings. Move the first person in expert group 1 to reporting group 1 the second person to reporting group 2, the third person to reporting group 3 and so on. Similarly the first person in expert group 2 moves to reporting group 1 the second person to reporting group 2 and so on.

The movement is shown in the diagram.


Whole group
o x
• *
o x
• *
o x
• *
o x
• *
Expert group1
Expert group 2
Expert group 3
Expert group 4
o o
o o
x x
x x
• •
• •
* *
* *
Reporting group 1
Reporting group 2
Reporting group 3
Reporting group 4

MODIFIED JIGSAW MOVEMENT
ADDITIONAL NOTE 1.1 (CONTINUED)

The movement is designed so that each reporting group has one member of each of the expert groups.

In the reporting groups each participant will report on the information collected in the expert group. In this way **everyone** has a turn as a reporter and everyone hears from all the other groups. It is important to let participants know at the start that they will all end up being reporters so they can keep good notes.

Note: If there are uneven numbers two people from one expert group may go to the same reporting group and share the reporting. Alternatively the leader may join a group to make up the numbers.

If there are four expert groups containing five members each then, after the movement, there will be five reporting groups containing four participants each. Similarly if there are three expert groups containing four members each the result will be four reporting groups containing three members each.

While the movement may appear daunting at first sight it is well worth persevering with because it has been shown in the trials to be a highlight with participants.

ADDITIONAL NOTE 1.2
AN EXAMPLE OF THE IMPORTANCE OF CONTEXT

A major project defined 'rurality' by correlating statistical indices against common sense definitions of rurality for several Australian areas.

The consequence was that the index of rurality was defined to be based on:

1 The number of vehicles per household: the more vehicles the more rural.

2 The proportion of households with TV sets: the less TVs the more rural.

3 The percentage of the work force in rural industries.

In the national context this definition works very well and provides a useful index of rurality. But when applied in Tasmania, a **rural** town— New Norfolk— became the most **urban** locality in the state!

New Norfolk is a mill town with very high proportions of the work force in manufacturing industry and very low proportions in rural industries. It has perfect TV reception and very high set ownership. Being a mill town with bus travel provided to the plant one car families are the norm. Yet in the Tasmanian context New Norfolk behaves as a rural centre in most ways.

Consequently the local context can make an index or procedure developed for a different, or broader, context quite unsatisfactory and could lead to invalid results being obtained.

ADDITIONAL NOTE 1.3
SUGGESTIONS FOR FOCUSING ON PARTICULAR EVALUATION QUESTIONS

Here is the task:

> An ALBE project management team wants to include 'everything' in the evaluation, but resources are limited. How would you go about encouraging them to focus on particular evaluation questions?

In response to the task participants might make some of the suggestions listed below.

1 Make a list of possible evaluation questions then prioritise them.

2 Ask a small number of key stakeholders to submit one question each.

3 Exclude all questions that are peripheral to the stated purpose of the evaluation.

4 Work out an estimated cost to tackle all the evaluation questions and compare with the funds available.

5 Ask for twice as much money and staff to conduct the evaluation as budgeted for.

6 Conduct a survey with stakeholders (e.g. using a nominal group technique—see Module 3) to determine which questions are essential for the evaluation.

Individuals will probably suggest a wide range of ideas.

MODULE 2

EVALUATION APPROACHES AND METHODOLOGIES

56 57

OVERVIEW OF MODULE 2

1 EVALUATION APPROACHES AND METHODOLOGIES

2 NOMINAL DURATION

Three hours

3 PURPOSE

This module provides an outline of the major evaluation methodologies and when they are used.

It introduces participants to evaluation for program development, for program clarification, for program improvement and evaluation for justification and accountability.

It gives details of these evaluation approaches and prepares participants to start using them for themselves.

The module introduces participants to the collection of performance information and performance indicators.

4 RELATIONSHIP TO COMPETENCY STANDARDS

See Item 8 of the Introduction, Relationship to Competency Standards.

5 PRE- OR CO-REQUISITES

Module 1 is highly recommended.

Section 1.10 is essential as participants must understand that different approaches to evaluation are adopted for different purposes.

6 SUMMARY OF CONTENT

An outline of the major evaluation methodologies and when they are used including evaluation for program development, for program clarification, for program improvement and evaluation for justification and accountability.

These evaluation approaches are described in detail and participants are prepared to start using them for themselves.

The module also covers the collection of performance information and performance indicators.

7 DELIVERY

Interactive workshop
Small and large group discussions
Group investigations with all participants reporting their findings to new groups

8 LEARNING OUTCOME

On completing this module participants will be able to complete the development of a plan for an evaluation project and document and justify the approach chosen.

9 ASSESSMENT CRITERIA:

On completing this module participants will be able to:

- differentiate between evaluation for
  - program development
  - program clarification
  - program improvement
  - program justification and accountability

  and explain how to use them appropriately
- initiate an evaluation designed to help the development of an ALBE program
- sketch the logic of a program to help with its clarification
- identify at least one approach to program improvement
- identify the information to collect for monitoring program performance
- develop performance indicators from performance information
- outline at least one approach to impact evaluation.

10 ASSESSMENT CONDITIONS

The presenter will be responsible for assessing

- participation and
- implementation of the plan for an evaluation project.

11 ASSESSMENT METHOD

Participants are assessed in two ways.

- They must show through discussion and other workshop activities that they understand the use of evaluation for different purposes.
- Their Participant's Evaluation Project, when completed, must show evidence that they have evaluated a program for an appropriate purpose; and that they can document and justify the approach chosen in an evaluation of their chosen ALBE program.

12 SUGGESTED RESOURCES

Lambert, F. & Owen, J. 1994, *A Guide to Program Evaluation,*, DEET

Workforce Literacy Training Package, Section 6, 'Guidelines for conducting literacy skills audits and training needs analysis', SA DETAFE & ACAL, 1993, pp. 115–132.

13 PRESENTER'S NOTE

This is a long module and it may be hard to complete it in 3 hours.

Presenters may wish to leave out units such as 2.4, 'Evaluation for program development', if participants are reasonably familiar with needs analysis and skills audits.

Unit 2.7, 'Evaluation for program justification and accountability', is essential to the course.

MODULE 2 CONTENTS

1 Welcome and agenda

2 Report on work to date in planning an evaluation

3 Key Aspect 2: Progressive focusing

4 Evaluation for program development

5 Evaluation for program clarification

6 Evaluation for program improvement

7 Evaluation for program justification and accountability

8 Participants' Evaluation Projects

9 Session evaluation and close

PG OHP H AN

CONTENT	PRESENTER'S GUIDELINES	RESOURCES
SESSION 2 (about 3 hours)		
2.1 **Welcome and agenda for the workshop** (10 minutes)	Welcome participants and outline the purpose of the module and the agenda for the session.	Transparency 2.1 'Agenda—Module 2'
2.2 **Report on work to date in planning an evaluation** (15 minutes)	Give participants the opportunity to report briefly on their progress and any difficulties, if they have decided to develop a plan for the evaluation project at their own workplace as a between-session task.	***Note:*** *Split up larger groups, appoint extra facilitators (if available) and move between the groups yourself.*
2.3 **Key Aspect 2: Progressive focusing** (5 minutes)	Remind participants: • that it is important to clarify the evaluation questions and • that it is usually impossible to evaluate the totality of a program. Refer back to progressive focusing. • It can usefully occur when developing the evaluation questions • It is prompted if the stakeholders have to – select the most important of the proposed evaluation questions – or prioritise them. • Focusing will also occur as data collection and data processing methodologies are being developed.	***Note:*** *Refer to Unit 1.9 re progressive focusing.* ◆ An example of the use of focusing is provided by the issues arising from **Case Study 1.**

PG OHP H AN

2.4 **Evaluation for program development** (15 minutes)	Recall Handout 1.3 which showed four purposes for evaluation. Then show Transparency 2.2. Note the timing, typical questions and typical decisions. As mentioned in Transparency 2.2 a typical decision to be made as part of evaluation for program development is: *'What needs should be given priority as a result of the program?'* In the ALBE context this often involves skills audits and needs analysis.	Transparency 2.2 'Evaluation for program development' (the first 2 columns of Handout 1.3)
	Display Transparency 2.3 which shows connections between literacy task analysis, literacy skills audit and literacy training needs analysis, leading to a curriculum response (the resulting program).	Transparency 2.3 'Training Needs Analysis'
	Remind participants that another example of evaluation for program development was the self-evaluation conducted in Unit 1.4. Ask participants to work in SMALL GROUPS (of about 4) to list aspects of a new workplace literacy program which could be examined in the early stages of its planning and implementation, in order to help with program development. Ask groups to share their findings with the WHOLE GROUP and record their suggestions on the whiteboard. ◆ Participants could use **Case Study 2** for further reading about an evaluation undertaken partly for program development.	

PG OHP H AN

2.5 **Evaluation for program clarification** (40 minutes)	Show Transparency 2.4 and explain that evaluation for program clarification is often concerned with clarifying the underlying logic of a program. Distribute two handouts you have prepared from the Guide: • Handout 2.1 which shows an example of a program logic statement developed for a workplace basic education program.; and • Handout 2.2 which gives a process for determining a program logic statement. Ask participants to form SMALL GROUPS to prepare an initial sketch of the program logic of the evaluation course they have experienced to date.	Transparency 2.4 'Evaluation for program clarification' Handout 2.1 'A Program Logic Statement for a Proposed Workplace Basic Education Program' (*The Guide*, fig. 4.2, p. 30) Handout 2.2 'A Program Clarification Framework' (*The Guide*, figure 4.3, p. 31)
	Arrange for groups to display sketches on the walls for individuals to examine during a break. Recommend the following for further reading, 'A Program Clarification Framework', *The Guide*, pp. 26–27. ◆ **Case Study 2** is an evaluation conducted in part for program clarification purposes.	An alternative activity would be for groups to sketch the logic of a program they are familiar with, such as developing an ALBE course.

2.6 **Evaluation for program improvement** (40 minutes)	Introduce the idea of evaluation to improve a program by focusing on its processes and delivery.	*See Additional Note 1.1 'Modified jigsaw movement'.*
	Present Transparency 2.5 showing the four topics: – implementation studies – site-level evaluation – refining developing programs – continuous improvement.	Transparency 2.5 'Approaches to program improvement'
	GROUP INVESTIGATION - JIGSAW MOVEMENT	
	Divide participants into FOUR EXPERT GROUPS. *(If there are less than 12 participants use 3 groups; ask one group to examine the first 2 topics.)*	
	Assign each group to one of the topics from Transparency 2.5 and give them the corresponding readings—Handout 2.3 (made from the Guide) and Handout 2.4. Ask the groups to find out about the assigned approaches from the readings and from their own experiences so they can each report on their findings to a new group.	Handout 2.3 'First three topics' (Guide, pp. 36–38) Handout 2.4 'Continuous improvement'
	Regroup into REPORTING GROUPS containing one member from each expert group. (Display Transparency 1.5 to show the movement.) Participants now report on their findings to the new groups.	Transparency 1.5 'Modified jigsaw movement'
	On returning to the WHOLE GROUP emphasise the importance of developing an evaluation culture, e.g. by integrating evaluation into ongoing program development when a program is being planned.	◆ See **Case Study 3** for an evaluation for program improvement.
	Note: *Useful references for further reading are given in the Guide, p. 41.*	

PG OHP H AN

2.7 **Evaluation for program justification and accountability** (40 minutes)	Show Transparency 2.6 to introduce evaluation for accountability purposes.	Transparency 2.6 'Evaluation for program justification and accountability'
	Outline the two approaches: – monitoring performance (Transparency 2.7) and – impact evaluation (Transparency 2.8).	Transparency 2.7 'Monitoring performance' Transparency 2.8 'Impact evaluation'
	MONITORING PERFORMANCE	
	Explain the use of performance monitoring as part of program management and planning cycles. (*The Guide*, p. 46)	
	Show Transparency 2.9 as a typical example. You might like to use the quote that such evaluation is: *'organisational methods of learning from experience'* (*The Guide*, p. 46).	Transparency 2.9 'Project management cycle'
	In SMALL GROUPS brainstorm all the possible performance information that could be collected about an ALBE program.	
	Stress the need to think widely at this stage and include both quantitative and qualitative information and information which might be difficult to collect or measure.	◆ **Case Studies 1 and 2** describe evaluations for accountability purposes which may be useful as additional reading.
	Share with WHOLE GROUP and list the suggestions on OHP or whiteboard. Continued next page	

PG OHP H AN

2.7 **Evaluation for program justification and accountability (continued)**	Show Transparency 2.10 which gives a definition of performance indicators.	Transparency 2.10 'Performance indicators'
	In the same SMALL GROUPS • decide which of the suggested items of performance information could be turned into a performance indicator • note any advantages or disadvantages.	
	Share results with the WHOLE GROUP.	
	Point out the validity and ethical issues in the use of performance indicators. (*Handout 2.5 has additional information.*)	Handout 2.5 'Validity and ethical issues' (*The Guide*, pp. 50–51)
GROUP INVESTIGATION–IMPACT EVALUATION	Show Transparency 2.11 to introduce impact evaluation as end-of-program evaluation of program outcomes.	Transparency 2.11 'Impact evaluation questions'
	◆ Mention **Case Studies 1 and 2** as examples of impact evaluations.	Transparency 1.5 for the jigsaw movement.
Note: *If time is pressing it may be necessary to leave out this activity or to allow only a short time for groups to meet.*	Divide participants into THREE EXPERT GROUPS and assign each group one of the three approaches to impact evaluation. Give each participant Handout 2.6 (made from the Guide).	Handout 2.6 'Client comparison evaluation, follow-up studies and modality tests' (*The Guide*, pp. 54–55)
	Form REPORTING GROUPS containing one person from each expert group. Participants report on the approach they studied.	

PG OHP H AN

2.8 **Participants' Evaluation Projects** (5 minutes)	Tel! participants, if they are undertaking an evaluation project, that the next step is: • to decide on the detailed evaluation approach they are going to adopt from the various methods outlined in this module. *If time permits they could work IN PAIRS to document the chosen evaluation approach during the session.* *Alternatively they could document it before the next session.*	
2.9 **Session evaluation and close** (10 minutes)	Ask participants to fill in the evaluation sheet before leaving and to hand it in. Encourage the WHOLE GROUP to critically appraise this different style of evaluation form, if time permits.	Handout 2.7 'Evaluation form 2'

OHP

AGENDA – MODULE 2

1 Welcome and agenda
2 Report on work to date in planning an evaluation
3 Key Aspect 2:
Progressive focusing
4 Evaluation for program development
5 Evaluation for program clarification
6 Evaluation for program improvement
7 Evaluation for program justification and accountability
8 Participants' Evaluation Projects
9 Session evaluation and close

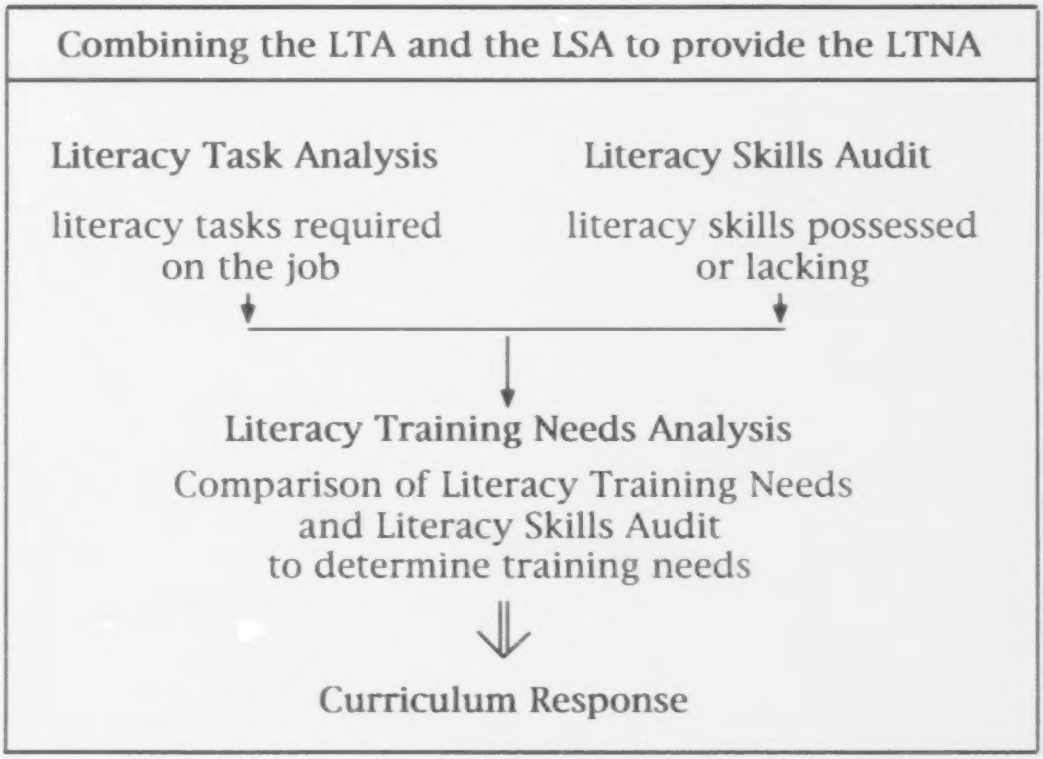
Transparency 2.1

EVALUATION FOR PROGRAM DEVELOPMENT

PURPOSES OF EVALUATION	**PROGRAM DEVELOPMENT** to make decisions about how best to develop a program to meet the needs, e.g. training needs analysis
FOCUS	context
TIMING	before
STAGE OF PROGRAM DEVELOPMENT	none
TYPICAL QUESTIONS	• What is the ideal or desired state of affairs? • What is the actual state of affairs? • Are there discrepancies between the desired and the actual? • What are the reasons for the discrepancies or needs?
TYPICAL DECISIONS	• Should the development of a program go ahead? • What needs should be given priority for action through an intervention or program?
ALBE EXAMPLES	*The Guide* pp. 19–20, pp. 22–23

Transparency 2.2

TRAINING NEEDS ANALYSIS



Combining the LTA and the LSA to provide the LTNA

Literacy Task Analysis

literacy tasks required on the job

Literacy Skills Audit

literacy skills possessed or lacking

Literacy Training Needs Analysis

Comparison of Literacy Training Needs
and Literacy Skills Audit
to determine training needs

Curriculum Response

from *Workforce Literacy Training Package*, p. 122

OHP

EVALUATION FOR PROGRAM CLARIFICATION

This form of evaluation is concerned with clarifying the underlying logic of a program. 'Logic' here is taken to mean 'a conceptual view of how a given program operates'.

It involves describing how a program is supposed to work, identifying cause-and-effect relationships and, in doing so, provides a rationale for the program as an intervention.

The Guide, p. 25

Transparency 2.4

APPROACHES TO PROGRAM IMPROVEMENT

- implementation studies
- refining developing programs
- site-level evaluation and action research
- continuous improvement

Transparency 2.5

EVALUATION FOR PROGRAM JUSTIFICATION AND ACCOUNTABILITY

A major purpose of evaluation is to account for the resources expended on the program under review.

Generally, accountability evaluation fulfils the need for someone directly associated with the program to report to somebody else, e.g. the funding agency.

In this scenario the manager justifies the funds and other resources used to provide the program.

Adapted from *The Guide*, p. 45

Transparency 2.6

MONITORING PERFORMANCE

Monitoring the performance of established programs is done to:

- provide information for the management of programs and
- provide the basis for external accountability.

It is an integral part of the management of programs. Managers need to:

- have information that will help them plan and control service delivery and
- account for their program and organisational responsibilities.

Adapted from *The Guide*, p. 45

Transparency 2.7

IMPACT EVALUATION

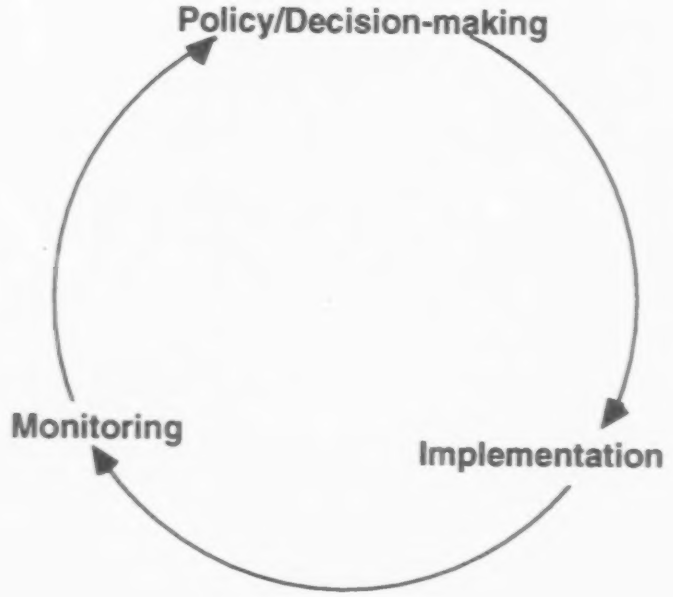
Impact evaluation also focuses on the delivery and the outcomes of well established programs, but the emphasis is on the assessment of outcomes. The primary purpose is to justify the worth of the program and to account for program expenditure.

These evaluations are usually conducted towards the end of a program.

Adapted from *The Guide*, p. 45

Transparency 2.8

PROJECT MANAGEMENT CYCLE



Policy/Decision-making

Implementation

Monitoring

Transparency 2.9

PERFORMANCE INDICATORS

A performance indicator is a simple statistic designed to give summary information which allows managers and others to make judgments about the effectiveness of programs.

Information is usually expressed as an index, rate or other numerical figure. Performance indicators are, in effect, substitutes for things that cannot be assessed directly.

For example, a medical symptom such as 'a raised temperature' is an indicator of illness.

NSW Public Service Board, 1986

Transparency 2.10

IMPACT EVALUATION QUESTIONS

What are the 'real goals' of the program and to what extent have they been attained?

What are the intended and unintended outcomes of the program?

How do differences in implementing the program affect the outcomes?

Does the program meet identifiable needs?

The Guide, p. 53

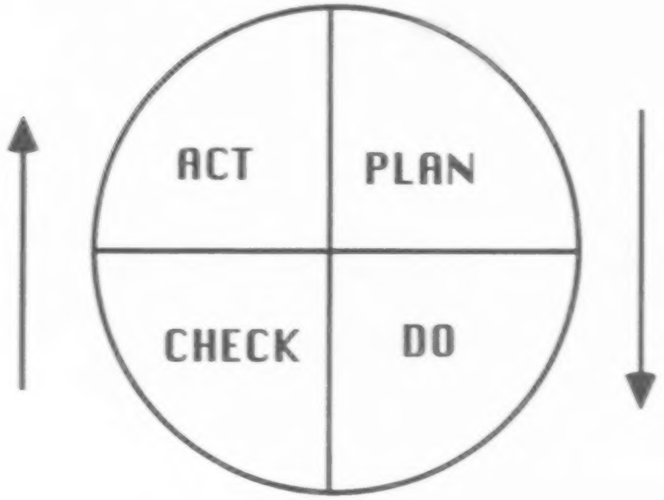
Transparency 2.11

CONTINUOUS IMPROVEMENT

Total Quality Management has popularised processes of continuous improvement suggested by W Edwards Deming and developed as Kaizen by Imai in Japan.

Continuous improvement is a small step improvement process. It usually involves a small group evaluating a process which they control to determine how it could be improved.

It involves a cyclic, four step process of planning, doing, checking and acting or standardising.



1 Plan

Identify the process or problem of interest. Collect information on the problem. Use the information collected to plan improvement. Determine what is to be done, who will do what, when, and ensure the required resources are available.

2 Do

Carry out the plan. Ensure stakeholders know about it. Provide any required training.

CONTINUOUS IMPROVEMENT (CONTINUED)

3 Check (or monitor)

Collect the necessary information to monitor what happens against the plan to ensure improvement is occurring. Discuss the results in the group and identify reasons for what is observed. Reflect on the findings and develop recommendations from the experience.

4 Act (standardise the improved performance)

If the checking stage confirms that improvements have resulted then make the new method the standard. Advise all who need to know about it and incorporate it in training for new staff. Establish further monitoring procedures to ensure the gains are maintained.

If the improvements are not achieved then review the data and decide whether to abandon the approach or to re-plan it in a different way and repeat the cycle.

Continuous improvement is very like other improvement cycles, for example action research involves a cycle of plan, implement, monitor and reflect.

Continuous improvement can be contrasted with the focus on innovative solutions popular in Australia.

EVALUATION FORM 2

Please give us your views of the following aspects of the session.

	V. good	good	fair	poor
1 The course content	☐	☐	☐	☐
2 The style of presentation	☐	☐	☐	☐
3 The venue	☐	☐	☐	☐
4 The organisation	☐	☐	☐	☐
5 The support materials	☐	☐	☐	☐
6 Group activities	☐	☐	☐	☐
7 Individual presentations	☐	☐	☐	☐
8 Getting what you came for	☐	☐	☐	☐

Other comments? ____

____

____

____

____

____

____

Handout 2.6

MODULE 3

INFORMATION COLLECTION ALTERNATIVES

97

96

OVERVIEW OF MODULE 3

1 INFORMATION COLLECTION ALTERNATIVES

2 NOMINAL DURATION

Three hours

3 PURPOSE

This module provides an outline of a large number of information collection methods which are widely used in evaluations. Participants are introduced to preparing and administering questionnaires.

4 RELATIONSHIP TO COMPETENCY STANDARDS

See Item 8 of the Introduction, Relationship to Competency Standards.

5 PRE- OR CO-REQUISITES

Modules 1 and 2 are recommended, but Module 3 can also be a stand alone unit for participants who understand the basic principles of evaluation.

6 SUMMARY OF CONTENT

A large number of information collection methods are described.
Additional information is provided about preparing and conducting questionnaires.

7 DELIVERY

Interactive workshop
Small and large group discussions
Group investigations with all participants reporting their findings to new groups

8 LEARNING OUTCOME

On completing this module participants will be able to choose data collection methods to be used for an evaluation project; and document and implement them.

9 ASSESSMENT CRITERIA

Participants will be able to:

- compare information collection alternatives, indicating the advantages and disadvantages of at least three methods
- design suitably worded questionnaires including the use of closed and open questions.

10 ASSESSMENT CONDITIONS

As for Module 2.

11 ASSESSMENT METHOD

Participants are assessed in two ways.

- They must show through discussion and other workshop activities that they understand different data collection alternatives and their purposes.
- Their Participant's Evaluation Project, when completed, must show evidence that they can employ appropriate information collection methods in an evaluation of an actual ALBE program .

12 SUGGESTED RESOURCES

A Guide to Program Evaluation, 1994, DEET

MODULE 3 CONTENTS

1 Welcome and agenda for the workshop
2 Report on work to date
3 Key Aspect 3: Making judgments
4 Collecting information
5 Information collection alternatives
6 Information collection: questionnaires
7 Participants' Evaluation Projects
8 Session evaluation and close

CONTENT	PRESENTER'S GUIDELINES	RESOURCES
SESSION 3 (about 3 hours)		
3.1 **Welcome and agenda for the workshop** (10 minutes)	Welcome participants. Outline the purpose of the module and the agenda for the session. Hand out the Session 3 evaluation form which asks participants to evaluate the session as it is delivered. *In Unit 3.7 they will be asked to discuss this approach to evaluation.*	Transparency 3.1 'Agenda —Module 3' Handout 3.1 'Evaluation form 3'
3.2 **Report on work to date** (20 minutes)	Ask participants to report briefly on their progress and any difficulties if they have been working on an evaluation as a between session task.	
3.3 **Key Aspect 3: Making judgments** (20 minutes)	Remind participants of the definition of evaluation which includes decision making: for decisions and judgments to have a validity beyond the views of an individual there needs to be a reference point. Show Transparency 3.2 with common referencing approaches: • norm referencing • competency or criterion referencing • standards referencing. Explain these and encourage brief discussion.	Transparency 1.3 'A definition of evaluation' Transparency 3.2 'Referencing evaluation judgments and decisions' *Additional Note 3.1 'Referencing evaluation judgments and decisions'*

PG OPT H AN

3.4 **Collecting information** (15 minutes)	Explain that having determined an evaluation question and methodology the next step is to work out how to collect the information that is needed.	
	Ask INDIVIDUALS to make some notes on the question (given in Transparency 3.3): *'If you wanted to determine how to improve a prison literacy program where would you get the information and how would you collect it?'*	Transparency 3.3 'Collecting information to improve a program'
	On returning to the WHOLE GROUP ask each individual to share: – one piece of information they would collect and – the method they would use to collect it.	
	Distribute Handout 3.2 with the advantages and disadvantages of information collection alternatives for future reference.	Handout 3.2 'Information collection alternatives' (*The Guide*, p. 9–11)
	◆ Further Reading: The information collection methods used in the three **Case Studies** could be identified.	

PG

3.5 **Information collection alternatives** (40 minutes)	*In this unit participants will explore in detail some particular techniques used to gather information.*	
GROUP INVESTIGATION – JIGSAW MOVEMENT	Place participants into EXPERT GROUPS to research two or three of the information collection techniques from Transparency 3.4.	Transparency 3.4 'Information collection alternatives'
	• If possible allow individuals to negotiate which methods to research.	
	• Exclude questionnaires as these will be covered separately.	
	• Provide groups with Handout 3.3. *The assessment of learning outcomes can be important for a particular evaluation but is not included here. A framework for describing outcomes for ALBE is given in ACTRAC 1993,* National Framework for Adult English Language, Literacy and Numeracy Competence, *ACTRAC Products, Frankston, Vic.*	Handout 3.3 'Information collection alternatives' (*The Guide*, pp. 91–114)
	Regroup into REPORTING GROUPS containing one member from each expert group. (See Transparency 1.5 and Additional Note 1.1.)	
	Ask participants to report on the techniques they have examined.	

PG [illegible] H AN

3.6 **Information collection: questionnaires** (40 minutes)	Ask participants to share with the WHOLE GROUP a few experiences they have had with filling in questionnaires. Remind the group of important features of using questionnaires. • Questionnaires are widely used because they can access a large number of people • They are cheap and convenient to administer. • Designing questionnaires requires considerable expertise. • It is important to: – be clear about the information needed – know the issues (a preliminary focus-group stage may be needed) – have a good design and – pilot the questionnaire, preferably with the designer present. QUESTIONNAIRE CHECKLIST Ask participants to work IN PAIRS to list possible problems with wording questions. Refer participants to the questionnaire checklist on pages 95 to 96 of Handout 3.3. Ask pairs to list the 5 most critical problems and to briefly share these with the WHOLE GROUP.	Handout 3.3 is also used for this unit. **Alternative activity:** Provide copies of a small questionnaire. Ask participants to use the checklist to critique the questionnaire and to suggest improvements.

3.6 Questionnaires (continued)	**Types of questions** Initiate a discussion IN PAIRS of the pros and cons of open and closed questions and allow brief reports to the WHOLE GROUP. Ask participants to work in SMALL GROUPS to design a small questionnaire to find out what colleagues at their workplaces know, or would like to know, about evaluation. Share questionnaires with the WHOLE GROUP.	*Page 111 of Handout 3.3 gives examples of open and closed questions and combination questions.*
3.7 Participants' Evaluation Project (15 minutes)	Tell participants who are working on an ongoing evaluation as part of the course that the next stage is data collection. Ask participants to • discuss IN PAIRS the data collection methods they might use for their own evaluation project and • record their thinking. Tell them to implement the data collection methods they have decided to use before the next session.	If participants are not conducting an evaluation of their own they could discuss the following: *'If you were conducting an evaluation of the services provided by an ALBE unit in order to help management make decisions about the unit's future what data collection techniques would you be likely to use?'*

PG OHP H AN

3.8 **Evaluation and close** (15 minutes)	Collect the evaluation sheets that participants have filled in during the session before they leave. Invite a WHOLE GROUP discussion of – the benefits and – the disadvantages of evaluating a training session using an evaluation sheet filled in during the session, rather than one distributed at the end.	

OHP

AGENDA – MODULE 3

1 Welcome and outline of the agenda for the workshop

2 Report on work to date

3 Key Aspect 3:
Making judgments

4 Collecting information

5 Information collection alternatives

6 Information collection — questionnaires

7 Participants' Evaluation Project

8 Session evaluation and close

Transparency 3.1

OHP

REFERENCING EVALUATION JUDGMENTS AND DECISIONS

norm referencing

competency or criterion referencing

standards referencing

Transparency 3.2

COLLECTING INFORMATION TO IMPROVE A PROGRAM

If you wanted to determine how to improve a prison literacy program where would you get the information and how would you collect it?

Transparency 3.3

INFORMATION COLLECTION ALTERNATIVES

1 Written responses, diaries etc.
2 Checklists or inventories
3 Rating scales, Delphi technique
4 Ranking scales and semantic differential scales
5 Written accounts
6 Observation forms
7 Face to face interviews and telephone interviews
8 Group interactions, focus groups etc.
9 Search conferences
10 Audio and video taping
11 Photography and unobtrusive techniques
12 Existing records

Transparency 3.4

EVALUATION FORM 3

Please make a note of your views about various aspects of the course as the session proceeds.

1 Things you liked

2 Things that could be improved

3 Your comments or suggestions

Handout 3.1

ADDITIONAL NOTE 3.1
REFERENCING EVALUATION JUDGMENTS AND DECISIONS

Evaluation involves making judgments about the program or activity being evaluated. Anybody can make judgments, and in most situations different people will make different judgments. Whose judgments will we accept?

For these reasons it is necessary to agree on some process which will reference, or validate, the judgments made as part of the evaluation. This is common in the ALBE field when assessing student performance: some sort of referencing system may be used to help ensure that valid and fair assessments are made.

The common referencing systems used in education are:

norm referencing
competency or criterion referencing
standards referencing

Norm referencing is a comparison to the performance of others, ideally to a large population of others whose performances are distributed on a normal curve. Tertiary entrance assessments are usually norm referenced.

Competency or criterion referencing makes judgments based on the competencies required to perform a particular task or the criteria for demonstrating a particular outcome. The judgments are referenced against the work or task.

Standards referencing makes judgments against some agreed standard of performance The standards are usually somewhat arbitrary but are either set by experts in the field or by agreement between most practitioners. The Australian Curriculum Profiles are standards for teachers to use in assessing student performance.

MODULE 4

COLLECTING AND ANALYSING EVALUATIVE INFORMATION

125

124

OVERVIEW OF MODULE 4

1 COLLECTING AND ANALYSING EVALUATIVE INFORMATION

2 NOMINAL DURATION

Three hours

3 PURPOSE

This module introduces participants to a number of important issues associated with collecting information for evaluation purposes, including sampling, reliability, validity and pilot studies.
The use of information analysing methods are outlined.

4 RELATIONSHIP TO COMPETENCY STANDARDS

See Item 8 of the Introduction, Relationship to Competency Standards.

5 PRE- OR CO-REQUISITES

Module 1 is recommended. Modules 2 and 3 would be helpful.
Module 4 can be used as a stand alone unit.

6 SUMMARY OF CONTENT

Issues associated with collecting information for evaluation purposes, including sampling, reliability, validity and pilot studies. Information analysing methods.

7 DELIVERY

Interactive workshop
Small and large group discussions
Group investigations

8 LEARNING OUTCOME

On completing this module participants will be able to implement a variety of techniques to analyse information during an evaluation project.

9 ASSESSMENT CRITERIA

Participants will be able to:

- list the reasons for undertaking pilot studies
- explain the use of random sampling and take simple representative samples
- make judgments about the reliability, validity and fairness of data collected for an evaluation
- explain the use of quantitative and qualitative data handling procedures.

10 ASSESSMENT CONDITIONS

As for Module 1.

11 ASSESSMENT METHOD

Participants are assessed in two ways.

- They must show through discussion and other workshop activities that they understand different information analysis techniques and how they are used.
- Their Participant's Evaluation Project, when completed, must show evidence that they can implement a variety of techniques to analyse information collected for an evaluation project.

12 SUGGESTED RESOURCES

A Guide to Program Evaluation, 1994, DEET.

MODULE 4 CONTENTS

1 Welcome and agenda for the workshop

2 Report on work to date

3 Reviewing information collection techniques

4 Pilot studies

5 Sampling

6 Key Aspect 4: Reliability and validity

7 Information analysis

8 Participants' Evaluation Projects

9 Evaluation and close

PG OHP H AN

CONTENT	PRESENTER'S GUIDELINES	RESOURCES
SESSION 4 (about 3 hours)		blank OHTs and pens for unit 4.7.
4.1 **Welcome and agenda for the workshop** (10 minutes)	Welcome participants and outline the purpose of the module and the agenda for the session.	Transparency 4.1 'Agenda—Module 4'
4.2 **Report on work to date** (15 minutes)	If participants have been working on an evaluation as a between session task they should be asked to report briefly on their progress and any difficulties.	
4.3 **Reviewing information collection techniques** (10 minutes)	Refer participants to the list on Handout 4.1 which provides a useful set of questions for reviewing the collection of data. Talk through the handout with the WHOLE GROUP and invite comments or questions.	Handout 4.1 'Questions for reviewing information collection techniques' (*The Guide*, p. 115)

4.4 **Pilot studies** (15 minutes)	Introduce the idea of a pilot study as a small scale practice run designed to test the effectiveness and efficiency of the data collection techniques and instruments. Use Transparency 4.2.	Transparency 4.2 'Pilot studies'
	Ask participants to work WITH A PARTNER to list possible reasons why it might be useful to undertake a pilot study before the actual evaluation data collection.	*More information about pilot studies is given in the Guide starting on page 115.*
	Share suggestions with the WHOLE GROUP and record them on the white board.	
4.5 **Sampling** (40 minutes)	Check that everyone understands the basic idea of sampling, i.e. 'collecting data from only part of a population rather than from everyone'.	
	Describe representative samples which are drawn in such a way that they accurately reflect the overall population.	
	Show Transparency 4.3 and allow questions and discussion.	Transparency 4.3 'Sampling'
	GROUP INVESTIGATION – JIGSAW MOVEMENT	
	Divide participants into FOUR EXPERT GROUPS to research the following different methods of sampling from Handout 4.2: – simple random sampling – stratified sampling – systematic sampling – purposive sampling	Handout 4.2 'Sampling' (*The Guide*, pp. 118–120)

PG OHP H AN

4.5 Sampling (continued)	Re-group into REPORTING GROUPS with one member of each expert group present to report on the sampling methods to the new group.	Transparency 1.5 shows the movement.
4.6 Key Aspect 4: Reliability and validity (40 minutes)	Show Transparency 4.4. Explain the meaning of 'reliable', 'valid' and 'fair'. An evaluation approach is **RELIABLE** if the same result is obtained on repeated occasions. An evaluation is considered **VALID** if it measures what it is intended to measure, i.e. it gives a true picture of the characteristic it is designed to measure. For example, if many respondents interpret a question in a different way to the designer of the questionnaire, the results will not be very valid. An evaluation approach should, of course, be **FAIR** to all those involved. Encourage participants to clarify their understanding of these terms by asking questions. ◆ The use of a variety of information gathering methods as a means of increasing the validity of data is mentioned in **Case Study 2**.	Transparency 4.4 'Reliability and validity' For more information see the Guide, pp. 122–123.

4.6 **Key aspect 4: Reliability and validity** (continued) (40 minutes)	GROUP INVESTIGATION Assign participants to THREE GROUPS : • one group considering reliability • one group considering validity • the third group considering fairness. Ask the groups to discuss this question: What precautions would they take in collecting data about the effectiveness of a literacy course conducted at a neighbourhood house to ensure – reliability? – validity? – fairness? Re-group into REPORTING GROUPS. Participants report on their findings to the new group.	*This unit uses the same movement as the previous section.*

4.7 **Information analysis** (30 minutes)	Have the WHOLE GROUP brainstorm some of the possible reasons why the data collected has to be analysed, e.g. – data reduction – data display – data communication to others – drawing conclusions – making meaning. Record brainstorm results on OHP.	blank OHTs and pens
	Assign participants to SMALL GROUPS of three or four.	Transparency 4.5 'Quantitative and qualitative data handling'
	Arrange for some groups to work on quantitative data and some on qualitative data. Use the following questions (as on Transparency 4.5). • Quantitative data handling: *'What can you do with numerical data about two different ALBE programs to allow the data to convey meaning to others?'* • Qualitative data handling: *'What could you do with descriptive and interview data about two different ALBE programs to allow the data to convey meaning to others?'*	*More detailed information on information analysis is given in the Guide on pp. 124–127, including useful references.*
	Groups choose a recorder/reporter and record their findings on OHT.	blank OHTs and pens
	Reporters share transparencies with the WHOLE GROUP.	
	Highlight any similarities or differences between quantitative and qualitative data handling after, or during, the reports.	

PG OHP H AN

4.8 Participants' Evaluation Projects (15 minutes)	Tell participants who are working on an ongoing evaluation as part of the course that the next stage is to analyse the information collected. Ask these participants to discuss IN PAIRS the techniques they might use to analyse their information and to record their intentions. Share some of these with the WHOLE GROUP, if time permits. Tell participants to analyse the data they have collected before the next session.	**Alternative activity:** Participants not undertaking an evaluation could work IN PAIRS to discuss how they would analyse the results which might be obtained using a ranking scale (Handout 4.4) with a large sample of adult education literacy teachers. For Handout 4.4 use the scale provided in 'Ranking Scale: an example' (*The Guide*, p. 109).
4.9 Evaluation and close (10 minutes)	Ask participants to fill in the evaluation sheet before leaving and to hand it in.	Handout 4.3 'Evaluation form 4'

AGENDA - MODULE 4

1 Welcome and outline of the agenda for the workshop

2 Report on work to date

3 Reviewing information collection techniques

4 Pilot studies

5 Sampling

6 Key Aspect 4:

Reliability and validity

7 Information analysis

8 Participants' Evaluation Projects

9 Session evaluation and close

Transparency 4.1

PILOT STUDIES

A pilot study is a small scale study designed to test the effectiveness and efficiency of selected data collection techniques and instruments. This involves trying out the instruments under conditions similar to those expected in the evaluation, and so it should include a range of people similar to those who will participate in the evaluation.

The Guide, p. 117

Transparency 4.2

SAMPLING

Sampling can provide an efficient and accurate way of obtaining information about a large number of cases. Just how efficient and accurate depends on the type of sample used, the size of the sample and the method of collecting data from the sample. In the end decisions about samples will be a compromise between cost, accuracy, the nature of the research problem and the art of the possible.

The Guide, p. 120

Transparency 4.3

RELIABILITY AND VALIDITY

An evaluation approach is considered **reliable** if the same result is obtained on repeated occasions.

An evaluation is considered **valid** if it measures what it is intended to measure.

An evaluation approach should, of course, be **fair** to all those involved.

Transparency 4.4

QUANTITATIVE AND QUALITATIVE DATA HANDLING

Quantitative data handling:

'What can you do with numerical data about two different ALBE programs to allow the data to convey meaning to others?'

Qualitative data handling:

'What could you do with descriptive and interview data about two different ALBE programs to allow the data to convey meaning to others?'

EVALUATION FORM 4

Please rate each of the statements below using the numbers 1 to 4 on the following scale:

Strongly agree	Agree	Disagree	Strongly disagree
1	2	3	4

1 Clear information was conveyed. ___

2 I could ask the questions I wished to. ___

3 I had the opportunity to work with a variety of people. ___

4 I couldn't understand the materials. ___

5 The session was well focused. ___

6 I enjoyed learning about evaluation. ___

7 The activities were varied. ___

8 We were rushed for time. ___

9 The presenter was a good role model. ___

10 I'm looking forward to the next session. ___

Any other comments?

____

____

____

____

Handout 4.3

MODULE 5

REPORTING AND ENSURING OUTCOMES

OVERVIEW OF MODULE 5

1 REPORTING AND ENSURING OUTCOMES

2 NOMINAL DURATION

Three hours

3 PURPOSE

This module introduces participants to the various ways of presenting and reporting on evaluation findings. Methods for ensuring effective outcomes from evaluations are developed collaboratively.

4 RELATIONSHIP TO COMPETENCY STANDARDS

See Item 8 of the Introduction, Relationship to Competency Standards.

5 PRE- OR CO-REQUISITES

Module 1 is recommended. Modules 2 to 4 would be helpful.
Module 5 can be used as a stand alone unit.

6 SUMMARY OF CONTENT

An introduction to

- various ways of presenting and reporting on evaluation findings and
- methods for ensuring effective outcomes from evaluations.

7 DELIVERY

Interactive workshop
Small and large group discussions
Group investigations
Focused problem solving

8 LEARNING OUTCOME

On completing this module participants will be able to plan the report of a specific evaluation project.

9 ASSESSMENT CRITERIA

Participants will be able to:

- identify how they would consider the needs of particular audiences when producing evaluation reports

- list
  - different ways of reporting on an evaluation and
  - key headings usually used in an evaluation report
- give reasons why it is desirable to build evaluation into a program from the start and prioritise ways to help an evaluation have impact.

10 ASSESSMENT CONDITIONS

As for Module 2

11 ASSESSMENT METHOD

Participants are assessed in two ways.

- They must show through discussion and other workshop activities that they know how to report on an evaluation project.
- Their Participant's Evaluation Project, when completed, must show evidence that they can lay out the plan of an evaluation project report in a standard format.

12 SUGGESTED RESOURCES

A Guide to Program Evaluation (1994) DEET.

MODULE 5 CONTENTS

1 Welcome and agenda for the workshop

2 Report on work to date

3 Evaluation as proof and evaluation as investigation

4 Determining audience needs

5 Reporting options

6 Report content

7 Key Aspect 5: Ensuring the impact of evaluation findings

8 Reporting on Participants' Evaluation Projects

9 Evaluation and close

PG OHP H AN

CONTENT	PRESENTER'S GUIDELINES	RESOURCES
SESSION 5 (about 3 hours)		Provide blank OHTs and pens for unit 5.4.
5.1 **Welcome and agenda for the workshop** (10 minutes)	Welcome participants, describe the purpose of the module and outline the agenda for the session.	Transparency 5.1 'Agenda—Module 5'
5.2 **Report on work to date** (20 minutes)	Ask participants who have been working on an evaluation to report briefly on their progress and any difficulties.	
5.3 **Evaluation as proof and evaluation as investigation** (30 minutes)	Ask participants to work IN PAIRS with Handout 5.1. Allow time for them to read it. ***Note:*** *This is a long extract and participants might skip over the more technical middle, sections.*	Handout 5.1 'Evaluation as proof and evaluation as investigation' (*The Guide*, pp. 121–124)
	Present the focus question (Transparency 5.2) for them to consider. Pairs report to the WHOLE GROUP.	Transparency 5.2 'Focus question'
	Note—Difficulty of Proof: *Groups will probably report that it may be difficult and costly to undertake true experimental designs in ALBE settings. Hence evaluation is unlikely to PROVE that particular programs are effective.*	◆ Note: Evaluation as proof is mentioned in **Case Study 2**.

5.4 **Determining audience needs** (25 minutes)	Ask pairs from the previous activity to join with another pair to form GROUPS OF FOUR. Provide blank OHTs and pens to record results. (Groups should appoint a recorder and presenter.)	blank overhead transparencies and pens
	Show Transparency 5.3.	Transparency 5.3 'Different audiences'
	Suggest to the groups that they might think in terms of an evaluation of the effectiveness of a staff training program and consider the needs of trainers, clients, staff managers and external funding agencies.	
	Reporters present suggestions to the WHOLE GROUP.	
5.5 **Reporting options** (15 minutes)	Ask INDIVIDUALS to • write down a few different ways in which the findings of an evaluation could be reported and • share some of these with the whole group, e.g. oral presentation, discussion session, written report.	
	Point out that Handout 5.2 lists several pairs of alternatives for reporting, e.g. written *versus* oral, progress *versus* final reports.	Handout 5.2 'Reporting options' (*The Guide*, pp. 131–133)
	Ask participants, in SMALL GROUPS or PAIRS, to list the benefits or disadvantages associated with the options for ONE such set of alternatives.	
	Report findings to the WHOLE GROUP.	

5.6 **Report content** (15 minutes)	Refer to Handout 5.2 which contains the content headings for a typical formal evaluation report. Ask the group to read the headings and suggest alternative ways to structure a formal report.	Handout 5.2 'Report headings' (*The Guide*, pp. 133–134)
5.7 **Key Aspect 5: Ensuring the impact of evaluation findings** (40 minutes)	Mention the concern that • evaluation findings should impact on the program concerned • many formal evaluation reports have little impact. *It is no use the evaluator learning about all the changes needed in a particular program if managers and other stakeholders remain unmoved.*	
	FOCUSED PROBLEM SOLVING *A full description of how to conduct a focused problem solving activity is contained in Additional Note 5.1.*	*Additional Note 5.1 'Conducting a focused problem solving activity'*
	Transparency 5.4 shows the different roles to be assigned.	Transparency 5.4 'Roles'
	Form SMALL GROUPS of four or five and assign – one person to be the problem owner – two to be questioners – one to be the facilitator and – a fifth (if available) to be timekeeper and monitor. Continued over	

PG OHP H AN

5.7 **Key Aspect 5: Ensuring the impact of evaluation findings** (continued) *Allow about 7 minutes for each questioning stage and about 10 minutes for the debriefing.*	Pose the following problems: *'How would you ensure that program decision-makers understand the results of evaluations and act on them? How could we incorporate effective change processes into an evaluation right from the start?'* (Transparency 5.5) Organise SMALL GROUPS to: – first undertake the three questioning stages – then debrief and – collect solutions ready to share with others. Ask group facilitators to report on solutions to the whole group. If different ideas emerge from the different groups, encourage further discussion, to see if a consensus emerges. You might mention – action research and – the continuous improvement process as approaches in which the program stakeholders and the evaluators are one and the same and hence are more able to implement findings.	Transparency 5. 5 'Focus Question—Problem Solving' *Additional Note 5.1 describes the process in detail.* ◆ Making private recommendations to management to assist the impact of an evaluation is mentioned in **Case Study 1**. ◆ The ability of participant evaluators to implement their findings is an issue arising in **Case Study 3**.

PG OHP H AN

5.8 **Reporting on the Participants' Evaluation Projects** (15 minutes)	At this point, give participants who are conducting their own evaluation, time to plan how they will report on their evaluation using the ideas developed in this session. Arrange for participants who want to, to work IN PAIRS to – discuss the reporting options they might use – select the most appropriate options and – make a plan of their evaluation reports. Advise them that they should prepare a draft of their evaluation report to share at the next session.	
5.9 **Evaluation and close** (10 minutes)	Ask participants to fill in the evaluation form before leaving and to hand it in.	Handout 5.3 'Evaluation form 5'

OHP

AGENDA – MODULE 5

1 Welcome and agenda for the workshop

2 Report on work to date

3 Evaluation as proof and evaluation as investigation

4 Determining audience needs

5 Reporting options

6 Report content

7 Key aspect 5: Ensuring the impact of evaluation findings

8 Reporting on Participants' Evaluation Projects

9 Session evaluation and close

Transparency 5.1

OHP

FOCUS QUESTION - EVALUATION AS PROOF OR EVALUATION AS INVESTIGATION?

In ALBE settings—

- when is:

  'evaluation as proof'
  most likely to be appropriate?

- and when is:

  'evaluation as investigation'
  most likely to be appropriate?

Transparency 5.2

OHP

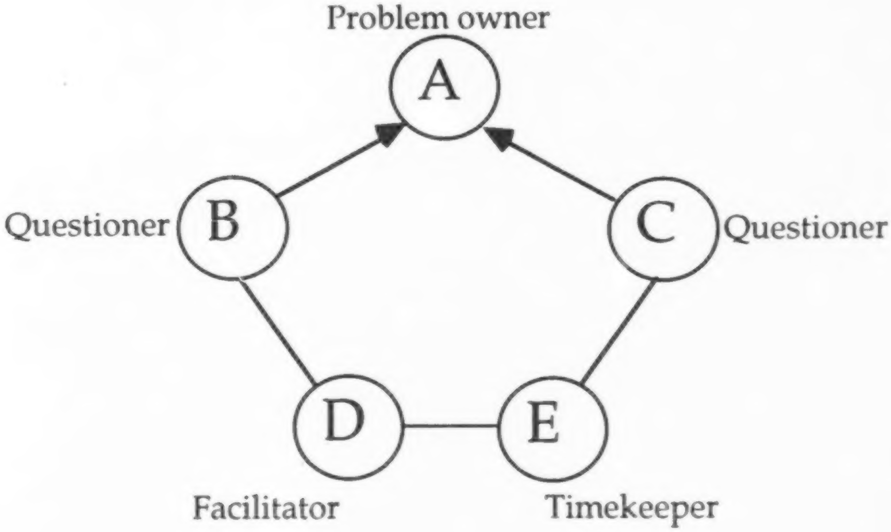
DIFFERENT AUDIENCES

Different audiences often require different information.

The Guide, p. 133

Transparency 5.3

ROLES


Problem owner
A
Questioner
B
C
Questioner
D
E
Facilitator
Timekeeper

Focused Problem Solving Roles:

A Problem owner
B Questioner
C Questioner
D Recorder, facilitator and reporter
E Timekeeper, question monitor and process observer

Transparency 5.4

OHP

FOCUS QUESTION – PROBLEM SOLVING

How would you ensure that program decision makers understand the results of evaluations and act on them?

How could we incorporate effective change processes into an evaluation right from the start?

Transparency 5.5

EVALUATION FORM 5

1. Please write down at least seven adjectives that describe your feelings about the session.

2 Circle the three adjectives (above) that you feel most strongly about.

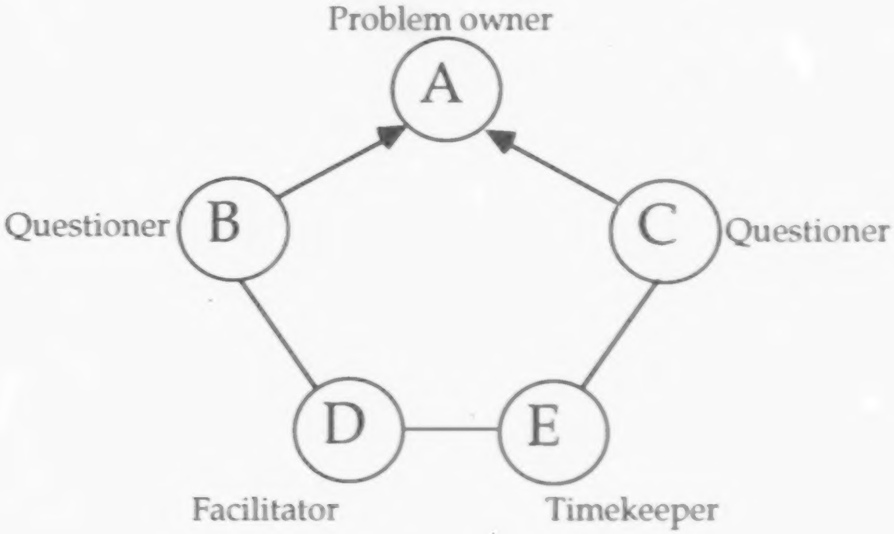
3 Use the circled adjectives in one or more sentences to provide constructive feedback to the presenter and course designers about the session.

4 Any other comments?

Handout 5.3

ADDITIONAL NOTE 5.1
FOCUSED PROBLEM SOLVING ACTIVITY

The activity is a group problem solving process for groups of four or five people.



Firstly a problem is identified and formulated as a question.

Next form groups of four or five participants and assign one person to be the problem owner (A), two to be questioners (B and C), one to be the facilitator (D) and a fifth (if available) to be timekeeper and monitor (E).

Group members offer to take on (or are assigned) the roles.

A The problem owner

The person who will adopt the problem as their own.

A can be any member of the group, not necessarily the one whose problem it is in the first place. This can offer a degree of objectivity to the problem solving process.

ADDITIONAL NOTE 5.1
FOCUSED PROBLEM SOLVING (CONTINUED)

B and C

The questioners

Two members, **B** and **C**, take on the role of asking questions only. They must strictly stick to the procedure of asking only one question at a time, in turn, and not to enter into any discussion.

About seven minutes are allocated to each of three question stages:

stage 1: fact seeking questions only,

stage 2: opinion seeking questions only, and

stage 3: solution directed questions only.

D Recorder, facilitator and reporter

D records all the possible solutions that emerge from the questioning stage on a vertical surface such as a white board, if possible, for all the group to see.

Once question time is over **D** facilitates a ten minute debriefing stage. All members share their thoughts feelings and opinions about the process. **D** then encourages further discussion of solutions and prepares the group's solutions ready to report to others.

E Timekeeper, question monitor and process observer

E is the timekeeper and monitors the nature of the questions asked. If they are not strictly of the type required **E** asks the questioner to rephrase the question appropriately.

E also takes the role of observer of the process in order to identify whether the initial problem question shifts during the activity.

Group facilitators report on solutions to the whole group.
If different ideas emerge from the different groups further discussion may be encouraged to see if a consensus emerges.

MODULE 6

GOOD PRACTICE IN EVALUATION

OVERVIEW OF MODULE 6

1 GOOD PRACTICE IN EVALUATION

2 NOMINAL DURATION

Three hours

3 PURPOSE

This module introduces participants to quality standards and good practice in evaluation. Issues arising from the management of evaluations and the use of consultants are outlined.

4 RELATIONSHIP TO COMPETENCY STANDARDS

See Item 8 of the Introduction, Relationship to Competency Standards.

5 PRE- OR CO-REQUISITES

Modules 1 and 2 are recommended. The other modules would be helpful.
Module 6 can be a stand alone unit for participants with a background knowledge of evaluation principles.

6 SUMMARY OF CONTENT

Quality standards and good practice in evaluation
The management of evaluations and the use of consultants

7 DELIVERY

Interactive workshop
Small and large group discussions
Group investigations

8 LEARNING OUTCOME

On completing this module participants will be able to present a sound report of an evaluation project:

9 ASSESSMENT CRITERIA

Participants will be able to:

- apply the principles of good practice in evaluation
- report on the priority order of standards for evaluations and describe some implications of the standards
- specify the rights of human subjects used in evaluations.

10 ASSESSMENT CONDITIONS

The presenter will be responsible for assessing
- participation and
- implementation of the evaluation project and the project report.

11 ASSESSMENT METHOD

Participants are assessed in two ways.
- They must show through discussion and other workshop activities that they understand how to
  - apply the principles of good practice in evaluation
  - use evaluation standards correctly and appropriately
  - respect the rights of subjects of investigations.
- Their Participant's Evaluation Project, when completed, must show that their understanding of the above issues has been put into practice in the report of their own program evaluation.

12 SUGGESTED RESOURCES

A Guide to Program Evaluation (1994) DEET.

MODULE 6 CONTENTS

1 Welcome and agenda for the workshop

2 Report on work to date

3 ALBE evaluation practice

4 Standards for evaluations

5 Key Aspect 6: The rights of human subjects

6 Managing evaluations and using consultants

7 Evaluation of the course and close

CONTENT	PRESENTER'S GUIDELINES	RESOURCES
SESSION 6 (about 3 hours)		Provide blank OHTs and pens for Unit 6.6.
6.1 **Welcome and agenda for the workshop** (10 minutes)	Welcome participants, describe the purpose of the module and outline the session content.	Transparency 6.1 'Agenda—Module 6'
6.2 **Report on work to date** (40 minutes)	Give participants who have been undertaking their own evaluation as a between-session task the opportunity to celebrate their production of a draft evaluation report.	
	Ask participants to work IN PAIRS to share reports and mentor each other's work using the following questions (Transparency 6.2): • What did you do well? • What did you learn? • What would you do differently next time?	Transparency 6.2 'Mentoring questions'
	Encourage members of the pairs to report briefly on each other's work to the WHOLE GROUP.	

6.3 **ALBE evaluation practice** (20 minutes)	Show Transparency 6.3 which lists some of the assumptions underlying acceptable practice for ALBE evaluations. Ask participants, working IN PAIRS, to outline some of the implications of these assumptions. For example, if you were conducting an evaluation in an ALBE context, what might be the meaning of 'be owned by the field'? Pairs report implications to the WHOLE GROUP.	Transparency 6.3 'Assumptions about acceptable evaluation practice for ALBE'
6.4 **Standards for evaluations** (40 minutes)	Mention the publication in the USA of a set of 30 standards for educational evaluation (*The Guide*, p. 141). Show Transparency 6.4 which has the four areas which encompass the standards, in order of priority. Ask INDIVIDUALS • to read the section from Handout 6.1 (p. 139, last paragraph of the first column) which explains the priority order of the standards and • to make notes. Ask participants to report their ideas to the WHOLE GROUP. Continued on next page	Transparency 6.4 'Priority areas for evaluation standards' Handout 6.1 'Standards for evaluations' (*The Guide*, pp. 139–141)

PG

6.4 **Standards for evaluations** (continued)	GROUP INVESTIGATION – JIGSAW MOVEMENT Put participants into FOUR EXPERT GROUPS and ask each group to research one of the following topics from Handout 6.1. – utility standards – propriety standards – feasibility standards – accuracy standards Ask participants to investigate the standards and suggest how they might assist the conduct of evaluations. Regroup into REPORTING GROUPS containing one member of each expert group who will report on the set of standards they studied. Return to the WHOLE GROUP and encourage the clarification of any unresolved issues.	*Additional note 1.1 has the detail about group investigations and Transparency 1.5 shows the movement.*
6.5 **Key Aspect 6: The rights of human subjects** (20 minutes)	Remind participants that one of the evaluation standards is that the rights of human subjects should be respected and protected. Discuss IN PAIRS: 'What are the rights of the humans involved in evaluations of ALBE programs?' Share findings in the WHOLE GROUP and record salient points on the whiteboard. Discuss how to advocate these standards in your own organisation.	

6.6 **Managing evaluations and using consultants** (35 minutes)	Comment that ALBE personnel, rather than conducting evaluations themselves, may be called on to – manage evaluations and – use evaluation consultants .	
	Divide participants into SMALL GROUPS to discuss the questions from Transparency 6.5, using a recorder and reporter.	Transparency 6.5 'Discussion question – the ideal consultant'
	Ask groups to sketch, with appropriate labels, the ideal evaluation consultant—as a joint task.	blank OHTs and pens
	Share overheads with the WHOLE GROUP.	
6.7 **Evaluation of the course and close** (20 minutes)	Provide a list of contact numbers and addresses of the participants, with their agreement, to assist with their networking and to offer mutual help with possible future evaluations.	
	Ask participants to fill in the end of course evaluation (Handout 6.2).	Handout 6.2 'Evaluation form 6'
	Explain that they may be contacted in a few weeks as part of the overall evaluation of this professional development course. Advise participants of the details if you plan to conduct a follow-up evaluation of their course at a later stage.	
	Bring the professional development course to a suitable close.	

REFERENCES

ACTRAC 1993, *National Framework for Adult English Language Literacy and Numeracy Competence,* ACTRAC Products, Frankston, Vic.

Bennett, B., Rolheiser-Bennett, C. & Stevahn, L. 1991, *Cooperative Learning: Where Heart Meets Mind,* Interactive Resource Book, Toronto.

Bomford, P. 1992, *Evaluation Research: Questionnaires and key documents from the evaluation of Workplace Basic Education Tasmania,* DEIRT, Hobart.

Bomford, P. 1992, *Workplace Basic Education Evaluation Report,* DEIRT, Hobart.

Grant, A. 1987, *Opportunity to Do Brilliantly: TAFE and the challenge of adult literacy provision in Australia,* AGPS, Canberra.

Impact of a National Framework for Professional Development of Adult Literacy and Basic Education Personnel 1994, TAFE National Staff Development Committee, Chadstone, Vic.

Lambert F. & Owen J. (in conjunction with S. Coates & J. McQueen) 1994, *Adult Literacy and Basic Education: A Guide to Program Evaluation,* DEET, Canberra.

SA DETAFE & ACAL 1993, Workforce *Literacy Training Package,* Adelaide.

Scheeres, H., Gonczi, A., Hager, P. & Morley-Warner, T. 1993, *The ABE Profession and Competence: Promoting Best Practice,* University of Technology, Sydney.

OHP

AGENDA – MODULE 6

1 Welcome and agenda for the workshop

2 Report on work to date

3 ALBE evaluation practice

4 Standards for evaluations

5 Key Aspect 6:
The rights of human subjects

6 Managing evaluations and using consultants

7 Course evaluation and close

Transparency 6.1

MENTORING QUESTIONS

Questions to help your partner reflect on his or her work:

- What did you do well?
- What did you learn?
- What would you do differently next time?

201

Transparency 6.2

ASSUMPTIONS ABOUT ACCEPTABLE EVALUATION PRACTICE FOR ALBE

ALBE evaluations must:

- be 'owned' by the field
- identify legitimate stakeholders (including the Commonwealth) and include them in decision making about evaluation
- make evaluation processes simple and accessible to stakeholders
- focus on critical areas of adult basic education, and
- be cost effective.

The Guide, p. 139

Transparency 6.3

PRIORITY AREAS FOR EVALUATION STANDARDS

The standards for evaluation of educational programs are grouped in the following four areas, in order of priority:

- utility
- feasibility
- propriety
- accuracy.

The Guide, p. 141

Transparency 6.4

DISCUSSION QUESTION –
THE IDEAL CONSULTANT

What criteria would you use for choosing a consultant for an evaluation?

What procedures would you put in place to help manage the evaluation?

Transparency 6.5

EVALUATION FORM 6

Please give your views of the following aspects of the overall course.

	V. good	**good**	**fair**	**poor**
1 The course content	☐	☐	☐	☐
2 The style of presentation	☐	☐	☐	☐
3 The support materials	☐	☐	☐	☐
4 Group activities	☐	☐	☐	☐
5 Getting what you wanted from the course	☐	☐	☐	☐
6 The usefulness of the course to your work	☐	☐	☐	☐

7 Can you suggest any improvements to the arrangement of the modules that make up the course?

____

____

8 Are there any additional materials or topics that you would like included in the course, or any you would like excluded?

____

____

____

9 List three things you have learnt from the course.

____

____

____

10 Other comments? (Please continue over the page if you wish.)

Handout 6.2

206

SUGGESTED READINGS

Bennett, B., Rolheiser-Bennett, C. & Stevahn, L. 1991, *Cooperative Learning: Where Heart Meets Mind*, Interactive Resource Book, Toronto.

Bomford, P. 1992, *Workplace Basic Education Evaluation Report*, DEIRT, Hobart.

Grant, A. 1987, *Opportunity to Do Brilliantly: TAFE and the challenge of adult literacy provision in Australia*, AGPS, Canberra.

Impact of a National Framework for Professional Development of Adult Literacy and Basic Education Personnel 1994, National Staff Development Committee, Chadstone, Vic.

Lambert, F. & Owen, J. (in conjunction with S. Coates & J. McQueen) 1994, *A Guide to Program Evaluation*, DEET, Canberra.

Handout 6.3

APPENDIX

CASE STUDIES

OF

ADULT LITERACY AND BASIC EDUCATION EVALUATIONS

208

INTRODUCTION

Three ALBE evaluations have been chosen as support materials to the course, *Professional Development for Program Evaluation: Evaluation for Adult Literacy and Basic Education Programs.*.

The evaluation reports have been condensed and reorganised as mini case studies showing how others have gone about evaluations, the methodologies they have adopted and the results and recommendations flowing from the studies. They range from national to local programs.

The three examples were chosen because, collectively, they demonstrate various significant features:

- evaluation used for a variety of different purposes
- processes which utilise different methodologies
- the use of uncomplicated data gathering and processing techniques.

Each one throws up interesting evaluation issues and should be relevant to a variety of ALBE staff.

SELECTED EVALUATIONS

The three evaluations selected were:

1 The **Workplace Basic Education Evaluation** conducted in 1991 by Pat Bomford of Department of Employment, Industrial Relations and Training (DEIRT now DIRVET), in Tasmania.

2 An independent evaluation of the **Impact of a National Framework for Professional Development of Adult Literacy and Basic Education Personnel**, conducted by Price Waterhouse Urwick in 1993 for the National Staff Development Committee for Vocational Education and Training (NSDC), then TAFE NSDC.

3 A study, **Opportunity to do Brilliantly: TAFE and the challenge of adult literacy provision in Australia**, conducted in 1984 by Audrey Grant for the Commonwealth Government and containing evaluations of two innovative adult literacy projects in Victoria.

HEADINGS FOR THE CASE STUDIES

Module 1 raises discussion of some of the components of an evaluation, the evaluation context, some headings for planning an evaluation, the evaluation question and the different purposes of evaluations. It includes discussions of stakeholders, audiences and why different evaluation purposes require different methodologies.

Accordingly, the evaluations included here as case studies have been summarised under the headings:

What is to be evaluated?
Context
The evaluation question
Purposes of the evaluation
Who is involved?
Audiences
Methodology

The following headings were also included:

Findings of the evaluation
Recommendations

In addition, any issues arising from the case studies are discussed.

CASE STUDY 1

WORKPLACE BASIC EDUCATION EVALUATION REPORT

WHAT IS BEING EVALUATED?

All 22 Workplace Basic Education (WPBE) courses conducted in Tasmania during 1991 and seven of the 26 courses completed during 1992. (Bomford, P. 1992, *Workplace Basic Education Evaluation Report*, DEIRT, Hobart and Bomford, P. 1992, *Evaluation Research. Questionnaires and key documents from the evaluation of Workplace Basic Education Tasmania*, DEIRT, Hobart.)

CONTEXT

The first workplace basic education courses started as a pilot program in Tasmania in 1988. The pilot was considered an outstanding success and the program was developed with a total of 36 courses conducted between 1988 and 1991. Courses were also run for Skillshare and the Commonwealth Rehabilitation Service, but these were not included in the evaluation which was restricted to courses conducted in workplaces.

Courses were initially government-funded, but evolved to become fee-for-service courses by the time of the evaluation.

From 1988 to 1991 WPBE in Tasmania was administered by a central unit in Hobart. In January 1992, WPBE was regionalised; the north-west, northern and southern regions of adult education became responsible for WPBE in their own regions.

THE EVALUATION QUESTION

The evaluation was designed to:

measure the effect of WPBE on course participants' work roles and personal development gauge the impact of WPBE on the work unit as a result of the individual undertaking the course

identify benefits and problems from the point of view of the service user and service provider

calculate the cost to the service user and service provider.

PURPOSE OF THE EVALUATION

The evaluation was for program justification and accountability purposes (see Unit 2.7 of Module 2). It includes both

- accountability aspects such as costs, benefits and effectiveness and
- impact issues such as evaluating the effect and impact of the courses.

WHO IS INVOLVED?

The stakeholders mentioned in the evaluation report were course tutors, managers of ALBE programs, managers and supervisors of businesses sponsoring courses, union representatives and course participants.

The evaluation was conducted by Pat Bomford, an in-house evaluator with the Evaluation and Research section of DEIRT (now DIRVET) in Tasmania.

AUDIENCES

Public service managers, industry, funding bodies.

METHODOLOGY

Minutes of the planning groups established at each workplace to advise on WPBE courses were examined to identify key issues.

Structured interviews were held with course participants, supervisors, union representatives and employer representatives (usually training officers). Some interviews were in person, some by phone. A detailed questionnaire was constructed for the interviews of each group and trialled in a separate work place. The interviews often deviated from the questionnaire and took the form of open discussion.

Course tutors prepared a report at the end of each course and a discussion group was held with six of the tutors.

FINDINGS

- Benefits from the WPBE courses for participants included:
  - participants' personal development
  - improvements in practical skills such as report writing and the use of calculators
  - gains in confidence and
  - improved communication skills.
- Teamwork in the workplace also improved.
- Twenty nine of the 52 participants interviewed said they learnt what they hoped to learn from the course while another 17 learnt 'to some extent'. Forty one of the 52 participants said that the course had helped them in life outside work.
- Employers were happy with the achievements of individuals on the courses. However, in workplaces where only a small number of participants completed courses employers found it difficult to justify the costs. Attendance problems occurred in five of the fourteen workplaces. Overtime, family commitments and shiftwork contributed to these.
- Employers could optimise course benefits and attendance by running courses totally within work time and cultivating a workplace culture that encourages training.
- Additional funding is needed to support WPBE. Employers were unwilling to spend more on courses. The author proposes that governments should share the financial responsibility with industry.
- Only 14% of course participants were female. Family commitments made women reluctant to participate in courses run outside of work time.
- WPBE needs strong and skilled marketing in the community and inside industry.
- There is a need for good communication between the regional structure for WPBE.
- The changes in industry, training and accreditation of training have important implications for WPBE.

RECOMMENDATIONS

A number of recommendations were made to WPBE in Tasmania, but they were not outlined in the report.

ISSUES ARISING

Focusing occurred at several stages. The questionnaires used in the interviews were refined in the light of experience. The plan to interview all 195 course participants had to be abandoned and interviews were only conducted with 52 participants. The plan to follow highly structured questionnaires in the interviews, which were even trialled before use, had to be abandoned when open discussion proved more useful in the interview situations. (Unit 2.3 of Module 2 has a discussion of progressive focusing.)

Impact or accountability evaluations (Unit 2.7 of Module 2) which try to compare costs and benefits are often difficult in social contexts. In this evaluation the financial costs were carefully calculated in dollar terms. The benefits were such things as personal development and report writing ability. It is very difficult to compare these and the report, wisely, makes no real attempt to do so.

Evaluations often lead to the identification of unexpected outcomes and issues: in this case, the low involvement of women in training programs.

The evaluator may have attempted to ensure the impact of the evaluation by directly, and privately, communicating the recommendations to management rather than indirectly through the report (see Unit 2.7 of Module 5). An in-house evaluator can sometimes make stronger recommendations privately to management than can be made in a public report.

CASE STUDY 2

IMPACT OF A NATIONAL FRAMEWORK FOR PROFESSIONAL DEVELOPMENT

WHAT IS BEING EVALUATED?

In 1991 the TAFE National Staff Development Committee (now NSDC) developed a strategic plan for professional development of ALBE staff called *A National Framework of Professional Development for Adult Literacy and Basic Education Personnel.*

The evaluation was to examine the extent to which the aims and objectives of the Framework were achieved in the first year of implementation. (*Impact of a Framework for Professional Development of Adult Literacy and Basic Education Personnel,* National Staff Development Committee for Vocational Education and Training, 1994, Chadstone Victoria.)

CONTEXT

Prior to the development of the Framework mentioned above little existed by the way of professional development for ALBE staff. There was a need for integrated structured professional development, with national applicability, and the consistent progressive development of skills and knowledge.

In 1991 the TAFE NSDC developed the Framework and began its implementation. Part of the implementation strategy was a summative evaluation of the Framework to be conducted at a national level by an independent evaluator.

THE EVALUATION QUESTION

To what extent have the aims and objectives of the National Framework been achieved in the first year?

PURPOSE OF THE EVALUATION

The evaluation purposes were:

> To identify any gaps in the Framework's objectives and delineate strengths and weaknesses of the professional development program. This is evaluation for program development (see Unit 2.4 of Module 2).
>
> To clarify, explain and interpret why any gaps exist. This is evaluation for program clarification (see Unit 2.5 of Module 2).
>
> To legitimate, account for and justify the resources used under the Framework. This is evaluation for program justification and accountability (see Unit 2.7 of Module 2).

It was also noted that:

> '...the evaluation report will allow the TNSDC to make decisions about the continued improvement of the PD program for ALBE staff.' (p. 5)

Hence the evaluation is also for program improvement (see Unit 2.6 of Module 2).

WHO IS INVOLVED?

The stakeholders are NSDC personnel, people conducting PD for NSDC, funding agencies, ALBE decision makers and ALBE staff.

The evaluation was conducted by consultants Price Waterhouse Urwick.

AUDIENCES

Commonwealth Government, NSDC, ALBE decision makers and ALBE staff

METHODOLOGY

The evaluation was conducted in four stages.

1. Analysis of existing information including the states' strategy documents and the formative evaluation questionnaires completed by PD participants at the end of training sessions.
2. Focus groups conducted in all states and territories with an average of four participants comprising PD managers, presenters of PD programs and selected program participants.

216

3 Interviews with ALBE managers in all states and territories.

4 A national survey sent to a sample of approximately 250 ALBE teachers in all states and territories.

FINDINGS

The findings of the evaluation were summarised against the objectives of the Framework.

Objective 1:
To provide a nationally coordinated approach to the planning, development, implementation and evaluation of ALBE professional development programs over the next two years.

The evaluators reported that while considerable progress has been made towards the achievement of the objective there was room for improvement.

Objective 2:
To ensure that adequate recruitment and induction processes are developed and implemented.

The evaluators found that this was the 'least well handled' of all the objectives because the selection and induction of staff are not handled through a national body.

Objectives 3 and 4:

3 *to promote opportunities for ALBE staff to develop additional competencies*
4 *to increase the competencies of ALBE personnel so that they can respond adequately to a broad range of learning situations or contexts.*

The finding was that these objectives were being achieved by the PD programs. However, there were some fundamental gaps, for example, the lack of a set of competency standards for ALBE staff and the fact that one third of the target population did not attend PD because of time or distance constraints.

Objective 5:
To ensure that current practice is informed by research, and theoretical implications are addressed through professional development opportunities.

The evaluators concluded that there was little doubt that this objective was being achieved by the PD activities.

RECOMMENDATIONS

The report makes 12 recommendations.

- Some are designed to help program development suggesting priorities for action, for example:

  Develop more specific programs based on competencies, such as interpersonal skills, computer skills, etc.

- Some are for program clarification suggesting how the program plan could be revised:

  Create national competency standards for ALBE staff.

  Standardise the format of the states' strategic and operational plans and define the performance indicators more precisely.

- Others are directed at program improvement, for example:

  Develop a strategy to improve access of programs to sessional and part-time staff.

  Develop all appropriate programs in flexible delivery format.

The different purposes overlap and are not separated in the report. In fact the report's findings and recommendations do not specifically refer to issues of program development and clarification, but concentrate on accountability issues and recommendations for the future.

ISSUES ARISING

The study illustrates how one evaluation can address all four common evaluation purposes outlined in Module 2.

The study very clearly identifies the purposes of the evaluation in a section (p. 5) headed, Functions (Why are we evaluating?). It would have been useful if these purposes were re-addressed in the findings and recommendations.

The use of four very different methods of information collection including interviews, focus groups and questionnaires improves the validity of the information collected (see Unit 2.6 of Module 4).

The evaluation does not 'prove' that the program is effective (see Unit 5.3 of Module 5), but the program is judged against its stated objectives.

CASE STUDY 3

OPPORTUNITY TO DO BRILLIANTLY:

TAFE AND THE CHALLENGE OF ADULT LITERACY PROVISION IN AUSTRALIA

WHAT IS BEING EVALUATED?

Inside a larger research project contributing to national policy development two innovative adult literacy projects were evaluated. They were:

1. The Council of Adult Education's literacy program in Melbourne in 1984 which provided small group literacy tuition mostly located in the city centre.
2. The Northern Districts Adult Literacy Group, known as the Glenroy Program, which, in 1984, used volunteer tutors for literacy teaching in the regional library and homes in the Broadmeadows area of Melbourne.

The complete research report was published in 1987. (Grant, A. 1987, *Opportunity to Do Brilliantly: TAFE and the challenge of adult literacy provision in Australia*, Australian Government Publishing Service, Canberra)

CONTEXT

The study was undertaken at a time of 'unprecedented Commonwealth government initiatives' in the area of adult literacy and basic education. The author comments:

> 'Never before has there been such groundswell and momentum at national, state and community levels, for righting the acute imbalance between the widespread need for adult literacy learning and the limited provision for actually meeting the demand.' (p. ix)

The relevant Commonwealth department was seeking guidance on the development of policy about adult literacy and on making funding recommendations to Government, so it commissioned the research project as part of the process.

THE EVALUATION QUESTION

The two evaluative case studies aim to describe what is distinctive about each program, what factors account for its successes and determine its applicability elsewhere, and what constraints or limitations are evident.

PURPOSE OF THE EVALUATION

The funding body sought a research study focused on one aspect of adult literacy which would contribute to the development of policy in the adult literacy area. As a project 'concerned with the interface between needs and provision' it can be seen as concerned with program improvement, but there is a sense in which it is research for program development of the new adult literacy programs funded after 1985.

The principal researcher conceived the research as an evaluation of two distinct models of adult literacy provision, reviewed in the wider context of recent developments in adult literacy. The evaluations were designed to describe:

> 'what is most distinctive about each program, what factors account for its success and determine its applicability elsewhere, and what constraints or limitations are evident.' (p. 2)

This is a style of evaluation for program improvement (see Unit 2.6 of Module 2), called responsive evaluation or naturalistic enquiry (see *The Guide*, p. 37 for more details). As a form of naturalistic enquiry it requires direct observation in the program setting and an openness to the issues which emerge.

Because the research team included four program tutor-coordinators who collected much of the interview data, the evaluations have quite a participant evaluation or action research style about them (see Unit 6 of Module 2 and *The Guide*, pp. 37–38). In these models evaluation for improvement is closely integrated into program delivery (*The Guide*, pp. 38-39).

WHO IS INVOLVED?

The stakeholders were ALBE staff, course participants and ALBE management. The chief researcher and author of the report was Dr Audrey Grant of La Trobe University. The research team included four adult basic education tutors/coordinators who collected much of the interview data.

AUDIENCES

Commonwealth government, funding agencies, ALBE managers and ALBE staff

METHODOLOGY

Case study methods were used to make a descriptive evaluation of the two models of adult literacy provision.

The case studies were based on interviews with a cross-section of program staff, tutors and students. Interviews were open-ended, responsive to the interviewees and based on common topics rather than a detailed schedule. Group discussions were held with the two program coordinators and the research team.

Records of program meetings, written observations and journals kept by participants, and coordinators' newsletters to tutors were also used as sources of information.
The data was analysed by the research team identifying major themes and recurrent patterns in the data. Attention was also given to factors unique to a particular program and factors accounting for or restricting the success of the programs.

FINDINGS

Detailed findings were reported about the individual projects.

Overall three major principles were identified:

1. that the multiplicity of needs and contexts for adult literacy and basic education points to the necessity for diversity of program and plurality of provision;
2. that a broad understanding of what becoming literate entails should inform all aspects of policy-making and implementation, program provision and evaluation research;
3. that the participatory-learning-support model of program provision pioneered in adult literacy be safeguarded and replicated to facilitate greater equity of access in adult education.

It is not possible here to summarise the 115 pages of the report describing the features of the two programs closely linked to the interview data. It can be said that

- the report brings to life the programs described and
- illuminates the issues they are grappling with.

It is clear from the report that the interviews and the feedback provided to tutors and coordinators influenced the evolution of the programs and individual teaching and hence contributed to program improvement. Recommendation 1 indicates that the researchers themselves are very clear about the value of this form of evaluation for program improvement.

In the spirit of the evaluation some reflective quotes from one of the four tutor-coordinator evaluators might be illuminative.

> 'It's really widened my vision of literacy. ....I think the actual interviewing of the students made me really understand how...literacy affects their lives or their lives have affected their literacy...The research, because people have opened up and talked about it, has given me in a very short time a picture of different people and how they work. I think that's made me look differently at myself...' (p. 159)

RECOMMENDATIONS

Six major recommendations were made.

1 The careful development of a continuing process of evaluation and participant research is an indispensable component throughout literacy provision: in ascertaining needs; in planning, developing and sustaining programs; in recruiting and training staff. Thereby that action-participant learning be the goal.

2 That programs be encouraged to develop flexible structures and hence new initiatives which better care for students' growth and adult workers' potential contribution. That additional adult literacy field officers be appointed to facilitate this development.

3 That policy implementation and allocation of funding promote devolution of responsibility so that decision making becomes increasingly the responsibility of experienced program participants. Such devolution must be preceded and accompanied by the development and funding of appropriate program support structures, staff development and training facilities, together with provision for participant-based evaluation and thereby for accountability.

4 That, in the proposed development of existing and planned future services, the quality of staff recruitment, training and development be seen as crucial.

5 That further research funding give priority to participant-evaluation, case-study based research. That this emphasis be promoted in commissioned research projects and in particular grants to institutions with an established commitment to teaching, postgraduate studies and qualitative research in language and literacy.

6 The guidelines for the allocation of funding apply to the above principles and recommendations and recognise that the long-term commitment to students necessitates guaranteed funding over a three-year period.

ISSUES ARISING

The evaluation illustrates how responsive evaluations can contribute to program improvement, particularly where the program staff also collect the evaluative information and hence get direct feedback on the results of their work (see Unit 2.6 of Module 2).

Such an approach also assists the implementation of evaluation findings (see Unit 2.7 of Module 5) because the evaluator and the action-taker are one and the same person.

The report shows how evaluations for program improvement, conducted on a small scale, can be used to contribute to principles and recommendations at the level of national policy-making and hence to program development (see Unit 2.4 of Module 2).

A limited case study methodology was adopted relying heavily on one data collection technique, interviews –a limitation forced by time and resource restrictions. This may have implications for the reliability, validity and fairness of the data (see Unit 2.6 of Module 4).

The use of participant-evaluators results in excellent understanding of contextual issues (see Unit 1.7 of Module 1), but may make it more difficult for the evaluators to make unbiased judgments.

ADULT LITERACY AND BASIC EDUCATION

A GUIDE TO PROGRAM EVALUATION

Produced by Dr. Faye Lambert & Dr. John Owen
in conjunction with
Sharon Coates & Julie McQueen

DEET

1990
INTERNATIONAL LITERACY YEAR

ADULT LITERACY AND BASIC EDUCATION

A GUIDE TO PROGRAM EVALUATION

Produced by Dr. Faye Lambert & Dr. John Owen
in conjunction with
Sharon Coates & Julie McQueen

This issue of ADULT LITERACY AND BASIC EDUCATION: A GUIDE TO PROGRAM EVALUATION has been specially printed by the National Staff Development Committee for Vocational Education and Training. It is provided free with the NSDC package, PROFESSIONAL DEVELOPMENT FOR PROGRAM EVALUATION: EVALUATION FOR ADULT LITERACY AND BASIC EDUCATION PROGRAMS, and must not be offered for sale.

© Commonwealth of Australia

This resource is Commonwealth copyright. It may be reproduced in whole or in part for study or training purposes subject to inclusion of an acknowledgment of the source; it is not to be copied for commercial use of sale. Reproduction for purposes other than those indicated above requires the prior written permission of the Department of Employment, Education and Training. Requests and enquiries concerning reproduction and copyright should be addressed to the Director, Literacy and ESL Section, Department of Employment, Education and Training, GPO Box 9880, Canberra City, ACT 2601.

FOREWORD

This publication is a product of a national project to develop good program evaluation practice within Adult Literacy and Basic Education (ALBE). The project has been managed by the Adult Community and Further Education Division of the Office of Training and Further Education in Victoria. It was funded as an International Literacy Year project from the Literacy and ESL section of the Department of Employment Education and Training (DEET). The views expressed herein do not necessarily represent the view of the Commonwealth Department of Employment, Education and Training.

The guidelines take into account consultations with a diverse range of stakeholders, the substantial experience of field-based contributors in conducting practical evaluations, and the conceptual understandings about program evaluation provided by the authors.

ABOUT THE PROCESS

In 1991 State and Territory delegates from the Australian Council for Adult Literacy surveyed or interviewed a range of people with an interest in the evaluation of adult literacy programs to gauge their knowledge and attitudes in relation to:

- goals/objectives of programs and whether or not these were evident to those involved
- the kinds of activities which would demonstrate the success or otherwise of programs
- instruments currently used to indicate success or otherwise
- ideas and opinions about evaluation instruments which could or should be developed to determine the achievement of program goals.

One of the key findings was that there was widespread confusion about the difference between student assessment and evaluation to the extent that, for many people interviewed, the terms seemed synonymous.

Subsequently, delegates came together with members of the project team and representatives of:

- State/Territory Systems
- National Training Board
- National Centre for Vocational Education Research
- Telecom Industries
- Workforce Literacy Providers
- Australian Manufacturing Council Secretariat
- DEET
- Confederation of Australian Industries
- National Board of Education, Employment and Training

to consider the findings and to discuss the interpretations to be placed on this information. Subsequently a report, which included the compilation of a list of evaluation models and case studies of evaluation practice was submitted from States and Territories.

During the next phase a thorough and comprehensive review of Australian and International literature on evaluation was conducted.

From the review and the report it was concluded that program evaluation should not be undertaken in isolation from a larger frame of reference which took into account program design and delivery. The concept of an interrelated and comprehensive system of evaluation at different levels of program provision was developed and tested in the field. It was important to understand the roles and functions of evaluation at all levels of operation, from the Commonwealth or national perspectives, through State/Territories and providers to the coal-face where teaching and learning takes place.

An important step in the final development of this publication was to develop frameworks which took into account the level of program operation and different types of evaluation (formative, summative, etc) which could be undertaken within ALBE.

Having verified these views through workshops and consultations, an expert evaluation team from the Centre for Program Evaluation at The University of Melbourne was commissioned to review these evaluation frameworks and to incorporate and extend them into an internally consistent set of evaluation guidelines. While primarily focussing on Adult Literacy and Basic Education uses, these guidelines may have wider applications.

The draft materials for this publication have been returned to the original data collectors and the Steering Committee for verification at various stages of their development. The authors acknowledge the support of ALBE personnel from across the country in the final production of these guidelines.

Project Manager

Sharon Coates	Adult Community and Further Education Division of the Office of Training and Further Education, Victoria. Responsible for project design and methodology, co-ordination and management, information dissemination, literature review, development of initial construct and review of documentation at all stages of the project.

Authors of the Guidelines

Dr Faye Lambert & Dr John Owen	Centre for Program Evaluation, The University of Melbourne.

Project Officers

Meryl Thompson	Adelaide College of TAFE; SA Responsibilities for collating and summarising research including evaluation practice in ALBE.
Julie McQueen	ACFE Division of the OTFE, Vic. Responsible for work on performance indicators and performance information.

Data Collection Team

ACAL delegates:	Rosa McKenna - Victoria Carlene Boucher- Executive Officer Marian Norton - QLD 1991 Helen Kebby - NSW Nancy Veal - ACT Sue Naylor - Victoria Margaret Carruthers - SA Jennie Bickmore-Brand - WA Di Kennedy - Tasmania Roy Pugh - Tasmania Julia Knight - NT
Steering Committee: Members:	Vanessa Elwell-Gavins - DEET Nicole Gilding - SA DEET Linda McLain - WA DEVET Marie Persson - ACAL then ANTA David Goldsworthy - ACFED, Vic Don Colless - NSW TAFE COM Rosa McKenna - NLLIA Ian Falk - Griffith University, QLD
Input and Advice:	Hugh Guthrie - NCVER Jim Thompson-DEVET

We would like to thank the following publishing houses for granting permission to reproduce material in this resource:

Allen & Unwin	NSW
Kluwer Academic Publishers	USA
Sage Publications, Inc.	USA

Finally, we wish to acknowledge the support of Ms Marion Brown, Administrative Assistant at the Centre for Program Evaluation, The University of Melbourne, who was responsible for assembling the material in the format in which it is presented in this publication.

TABLE OF CONTENTS

Chapters

1 Developing an Evaluation Culture for ALBE 1
2. The Three Stages of Evaluation 5
3. Evaluation for Program Development 17
4. Evaluation for Program Clarification 25
5 Evaluation for Program Improvement 35
6. Evaluation for Program Justification/Accountability 45
7. Information Collection Alternatives 93
8. Collecting and Analysing Evaluative Information 117
9. Reporting 133
10. Good Practice in Evaluation 139

Figures

2.1 Planning (Stage One) An Overview of Evaluation Purposes 8
2.2 (Stage Two) Information Collection Alternatives 9
2.3 Collecting and Analysing Evaluative Information (Stage Two) 12
2.4 Reporting (Stage Three) 14
2.5 An Evaluation Planner 15

3.1 Evaluation for Program Development - Questions to Ask 22
3.2 An Evaluation for Development Framework 23

4.1 Program Clarification - Questions to Ask 29
4.2 A Program Logic Statement for a Proposed Workplace Basic Education Program 30
4.3 A Program Clarification Framework 31
4.4 Reviewing Program Plausibility 32
4.5 An Objectives Hierarchy for a Proposed Workplace Education Program 33
4.6 An Expanded Objectives Hierarchy for a Proposed Workplace Basic Education Program 34

5.1 Program Improvement - Questions to Ask 42
5.2 Program Components: Reviewing Service Provision 43

6.1 Monitoring Evaluation - Questions to Ask 57
6.2 Project Management Cycle 58
6.3 An Objectives Hierarchy for a Proposed Workplace Education Program 59
6.4 An Expanded Objectives Hierarchy for a Proposed Workplace Basic Education Program 60
6.5 Post-Program Monitoring system (PPM) Labour Market Programs (DEET) 61
6.6 Natmiss Information Flow 62
6.7 A Structured Approach to the Development and Implementation of Performance Indicators 63

6.8-6.12 - Level of Operation: National. Sample performance indicators:

6.8 Policy and Planning 65
6.9 Program Development and Implementation 67
6.10 Personnel Management and Professional Development 68
6.11 Curriculum 69
6.12 Information Services- Financial and Resource Management 70

6.13-6.17 - Level of Operation: State/Territory. Sample performance indicators:

6.13 Policy and Planning 71
6.14 Program Development and Implementation 72
6.15 Personnel Management and Professional Development 73
6.16 Curriculum 74
6.17 Information Services and Financial/Resource Management 75

6.18-6.22 - Level of Operation: Region. Sample performance indicators:

6.18 Policy and Planning 76
6.19 Program Development and Implementation 77
6.20 Personnel Management and Professional Development 78
6.21 Curriculum 79
6.22 Information Services and Financial/Resource Management 80

6.23-6.27 - Level of Operation: Provider. Sample performance indicators:

6.23 Policy and Planning 82
6.24 Program Development and Implementation 83
6.25 Personnel Management and Professional Development 84
6.26 Curriculum 86
6.27 Information Services and Financial/Resource Management 89
6.28 Impact Evaluation - Questions to Ask 90

7.1 Rating Scales: Examples 109
7.2 Guidelines for Conducting a Delphi Study 110
7.3 Ranking Scale: An Example 111
7.4 Semantic Differential Format: An Example 112
7.5 Different Response Formats: An Overview 113
7.6 Guidelines for Questionnaire Layout 114
7.7 Five Dimensions along which Fieldwork Varies 115
7.8 Different Types of Evaluation Research Interviews: Characteristics, Strengths, Weaknesses 116

8.1 Purposeful Sampling Strategies 130
8.2 Matrix Display 131

Chapter 1 DEVELOPING AN EVALUATION CULTURE FOR ADULT LITERACY AND BASIC EDUCATION (ALBE)

What is meant by 'evaluation'?

In its broadest context, **evaluation is the collection and analysis of information in order to facilitate informed decision making.** The objects of evaluation include planning, programs, policies, or organisations.

Why the need for widespread understanding of evaluation within the Adult Literacy and Basic Education (ALBE) field?

Literacy is fundamental to economic development and social justice. No longer is it a peripheral concern but rather an important corollary of labour market programs and economic restructuring (LoBianco, 1989). The realisation that literacy is 'context-dependent' and that the rate of change may be outstripping the literacy levels of the nation has led to a strong imperative to maximise the effectiveness and efficiency of programs in meeting the increasingly complex and specific needs of those requiring literacy and basic education.

In a climate of decreasing public sector resources, 'evaluation continues to be increasingly used as a reporting and accountability strategem by funders and other authorities. Yet their purposes of providing value for money are not served well if the evaluation does not 'get at' either the qualitative nature of people's experience, or if it counts the wrong things, or at the wrong times, or if it is done for the wrong reasons' (Kenny, 1991).

There is the need not only for improved methods of evaluating at all levels of ALBE operations but also for a recognition that evaluation is part of a continuous feedback process. Administrators and providers must devote time and resources to implement findings arising from evaluation efforts.

Historically, the development of an evaluation culture within ALBE operations has been hampered by:

- a tendency to see evaluation as 'student assessment'
- a belief that evaluation is time-consuming, complex and difficult and something different from the evaluative tasks already built into existing practices
- a high value on individualism amongst providers
- a lack of useful/timely feedback from the range of program data collected
- a lack of policy, planning, infrastructure and resourcing support
- fear of the effects of evaluation on future funding.

These guidelines represent one of a range of projects initiated to assist policy makers, administrators and providers address these factors. Another, for example, is the National Framework of Adult English Language, Literacy and Numeracy Competence which has been designed to guide and inform the development and review of more specific programs for learning in English language, literacy and numeracy. The National Framework will be used for program planning and curriculum development, as well as for curriculum accreditation and evaluation.

The Evaluation Guide

a. Purposes

To improve the user's general understanding of the range of approaches and methods used to achieve different evaluation purposes.

To provide practical advice that will assist those at all levels within ALBE operations to undertake evaluations that will lead to improved efficiency and effectiveness.

To point to ways of making evaluation techniques an integral part of program administration and implementation.

To demonstrate the application of evaluation processes within the field of adult literacy and basic education, and in particular, focus on the selection and use of performance indicators reflecting the operational functions at national, state/territory, regional and provider level.

b. **Audiences**

Systems, departments, organisations and individuals operating at national, state, regional and local levels.

c. **Structure**

The material contained in these guidelines represents the latest conceptualisations about the practice and theory of program evaluation. Wherever possible, these ideas have been related to the ALBE area.

Chapter Two provides an introduction to the three fundamental stages of evaluation:
•planning
•collecting and analysing information
•reporting the findings.

The chapter provides a set of overview frameworks which summarise key ideas about each of the three stages.

The choice of the most appropriate form of evaluation approach is essential if the evaluation findings are to have an impact. Chapters Three to Six contains basic advice about each of the major evaluation forms to help the reader select the most appropriate one(s)..

Chapters Seven and Eight deal with data collection and analysis. Issues discussed include a range of data collection instruments, ideas about sampling and guidelines for data analysis.

Chapter Nine discusses evaluation reporting and provides some viable alternatives for ensuring that the findings are used by evaluation stakeholders

Finally, Chapter Ten discusses issues associated with general evaluation practice, including; guidelines for evaluations conducted for and/or by insiders, determining priorities for evaluation, general principles for maximising the utility of evaluation findings, and ethics of evaluation.

While the reader is frequently referred back to the overview frameworks given in Chapter 2, all other figures related to the text within each chapter are provided at the end of that chapter. Further readings in relevant evaluation topics are listed under the heading of references. In addition to these references, a bibliography is included which provides a more general reading list on evaluation with an emphasis on adult education and literacy.

Some users of these guidelines may feel comfortable using the material without assistance. However, some may need some professional development activities to assist them to use the guidelines to their full capacity. It is our understanding that such support will be forthcoming in the near future from ALBE sources.

References

LoBianco, J. (1989) Foreword in 'No Single Measure, A Survey of Australian Adult Literacy, Wickert, R. Institute of Technical and Adult Teacher Education, Sydney College of Advanced Education.

Kenny, S. (1991) Foreword in 'Everyday Evaluation On The Run', Wadsworth, Y., Action Research Issues Association (Incorporated), Melbourne.

Stage One:	Planning
Stage Two:	Collecting and Analysing
Stage Three:	Reporting

Decisions and plans made in each of these stages are interdependent and should be woven into an overall evaluation design. An **evaluation planner** provides a framework for this and is given on page 15.

The **PLANNING** framework in Figure 2.1 is designed to give conceptual and practical guidance in determining the most appropriate approach to evaluation in a given situation. It is based on the assumption that evaluation strategies can not only be used to justify the worth of an existing program, but that evaluation also encompasses activities likely to aid the development of programs and to increase an understanding of how they should operate. Five basic purposes for evaluation can be identified:

- Program Development
- Program Clarification
- Program Improvement
- Program Monitoring
- Program Justification

Note: The term 'program' here is used in the generic sense to encompass planning, policies and programs. Hence, 'program evaluation' could be substituted by 'planning evaluation' or 'policy evaluation'.

The **focus** of these different forms of evaluation varies from the context or circumstances within which a program might be developed and implemented, to a focus on the plan or design of a program, the actual delivery of the program or both the delivery and the outcomes of the program. They are therefore concerned with programs at different **stages of development** ranging from a situation where no program exists at all to where a program has been operating for a period of time and could be considered as 'established'.

The overview identifies questions and decisions typically associated with each of these forms of evaluation. Having identified which best describes the evaluation task being attempted, reference can be made to it in more detail in subsequent sections identified by page numbers in the overview diagram. However, it is important to be aware that real-life evaluations frequently range across the evaluation forms presented here.

Below are some evaluation scenarios which might well arise within ALBE operations. Even at this early stage and without reading about each form of evaluation in detail, the reader may wish to see if they can classify each of these scenarios into one of the forms, drawing upon the criteria of purpose, focus, stage of development and timing identified in the overview.

a. A tutor at a neighbourhood house is concerned about the ways she uses the limited amount of time she has with her adult classes in the evenings. She asks a colleague to observe the classes and assist in collecting information about her activities.

b. A State Vocational Education and Training agency has called for some research on communication requirements in the workplace to assist in developing policies and plans for literacy programs in industry.

c. Regional co-ordinators of ALBE would like all providers to adopt some standard procedures with respect to enrolling students and placing them on courses. While these procedures are largely in place, it is felt desirable that they should be made explicit in the context of an overall program plan.

d. State providers, responding to Commonwealth funding agencies, developed a set of performance indicators to monitor the effects of their ALBE operations.

e. A TAFE administrator has some concerns about the effectiveness of the college's provision for adult literacy and basic education and decides to initiate a comprehensive evaluation of the impact of the program.

f. A Regional Co-ordinator is concerned that there are 'pockets of need' in her area that are not being served by existing providers. She requests that a survey of the local community be undertaken to address this concern.

Having decided on the basic evaluation purpose, a plan can be developed for **COLLECTING AND ANALYSING** the information required.

Figures 2.2 and 2.3 provide a guide for making decisions about information collection. This framework is divided into two parts:

- Figure 2.2 identifies the range of information collection alternatives that might be appropriate for any particular evaluation, highlighting some of the potential advantages and disadvantages associated with their use in different situations.

- Figure 2.3 highlights the importance of selecting information collection methods and instruments that best serve the **key** issues identified in the evaluation and that can be used given the available resources and the time-frame for decision-making. Frequently, both qualitative and quantitative data are used to respond to specific evaluation questions and ideally more than one source of information is used to address a particular issue. The framework recommends the systematic review of information collection techniques and the testing of these techniques via a pilot study when possible. Where the evaluation seeks to collect information from different sites and/or a large group of people, it may be impractical to include them all. Taking a sample from within these 'total populations' enables the study to be kept manageable and within budget. Hence, alternative sampling techniques become important. For some evaluation studies, valid interpretation of the findings will require the setting up of controls to provide a basis for comparison. Part of the planning in the information collection stage involves anticipating the strategies to be used in analysing the data and some commonly used information analysis tasks are listed. Finally, of the standards for evaluation developed by the Joint Committee on Standards for Educational Evaluation (1981), those which apply in particular to information collection are listed.

Figure 2.4 discusses **REPORTING**. A common expectation of evaluation studies is that the end-result will be a substantial written document. However, concern about the lack of use of evaluation findings has led to the development of a vastly expanded repertoire of reporting strategies which attempt to respond directly to the needs of stakeholders (those with a legitimate interest in the program). It is important to note that the focus here is on the reporting of evaluation findings. This is not the same as reporting the results of student assessment, say in the areas of literacy and numeracy competence. As mentioned in the Introduction, evaluation is not the same as student assessment, although assessment information might be used to decide on the effectiveness of a given program.

Figure 2.4 identifies the range of reporting options commonly used, possible communication forms for different audiences, and the standards for evaluation which relate most directly to the reporting process.

Figure 2.5 is **AN EVALUATION PLANNER**, which provides a format for putting together the three stages into a coherent evaluation.

References

de Vaus, D. A. (1990) Surveys in Social Research. (2nd edition). Allen and Unwin, North Sydney, NSW.

Kosecoff, J. & Fink, A. (1982) Evaluation Basics A Practitioner's Manual, Sage Publications, London, New Delhi.

Morris, L., Fitzgibbon, C., & Freeman, M., (1987) How to Communicate Evaluation Findings, Sage Publications, California.

Owen, J.M. (1993) Program Evaluation Forms and Approaches, Allen and Unwin, St.Leonards, NSW Australia.

FIGURE 2.1: PLANNING (STAGE ONE) AN OVERVIEW OF EVALUATION PURPOSES

What to Evaluate? planning, programs, policies and organisations

PURPOSES OF EVALUATION Note: The term program is used here in the generic sense to encompass planning, policies and programs.	**PROGRAM DEVELOPMENT** to make decisions about how best to develop a program to meet the needs e.g. training needs analysis.	**PROGRAM CLARIFICATION** to clarify the logic of a program to assist in the effective delivery of outcomes.	**PROGRAM IMPROVEMENT** to improve the implementation of a program.	**PROGRAM JUSTIFICATION/ACCOUNTABILITY** to monitor program delivery and outcomes, to determine program worth.	
				MONITORING EVALUATION (accountability)	IMPACT EVALUATION (justification of worth)
FOCUS	context	design	delivery	outcomes/delivery	outcomes/delivery
TIMING	before	during	during	during	after
STAGE OF PROGRAM DEVELOPMENT	none	development	development	established	established
TYPICAL QUESTIONS	• What is the ideal or desired state of affairs? • What is the actual state of affairs? • Are there discrepancies between the desired and the actual? • What are the reasons for the discrepancies or needs?	• What are the intended outcomes of the program and how is the program designed to achieve them? • What program elements need to be changed to maximise its potential to achieve the intended outcomes? • What aspects of the program should be considered in a program evaluation? • Which of these aspects are amenable to evaluation? • What aspects require further attention before being evaluated at a later date?	• How is this service or activity going? • Is it working? • How is it affecting the target group or clients? • What specific aspects need improvement?	• Is the program reaching the target population? • Is it being implemented in the ways specified? • Is it effective? • How much does it cost? • What is the cost relative to effectiveness?	• What are the 'real goals' of the program and to what extent are they attained? • What are the intended and unintended outcomes of the program? • How do differences in implementing the program affect the outcomes? • Does the program meet identifiable needs?
TYPICAL DECISIONS	• Should the development of a program go ahead? • What needs should be given priority for action through an intervention or program?	• Is there a need to update or revise the plan or logic of the program currently being delivered?	• Should there be changes in the way the program is delivered? • Should policy associated with the program be reviewed or updated? • Should the program be modified so that it is delivered more efficiently?	• Is the funding allocated to the program justified in terms of what criteria? • Can the public have confidence in the services offered through the program? • Can praise or blame for the success or failure of the program be allocated to program personnel? • Should the program be discontinued because of its failure to provide adequate levels of outcomes?	
ABLE EXAMPLES	pp. 19-20, pp.22-23	pp.29-30, pp.33-34.	pp.38-40, pp.42-43.	p.57, pp.59-60, pp.63-90	

BEST COPY AVAILABLE

FIGURE 2.2: (STAGE TWO) INFORMATION COLLECTION ALTERNATIVES

(This section draws from the work of Kosecoff & Fink, 1982: 117-119, and Owen, 1993: 56-57.)

	Advantages	Disadvantages
1. Direct from Individuals p.93-100		
A. self-reports		
a. written responses, diaries or anecdotal records	Allows people to describe unique situations in their own words.	Often not well completed or maintained. Quality highly variable. Difficult to score and interpret.
b. checklists or inventories	Easily completed and scored.	
c. rating scales eg. Delphi techniques	Easily completed by respondents. Facilitates analysis of judgemental data.	Responses potentially biased by different rating standards of individuals. Response is limited by rating categories.
d. ranking scales	Easily completed and quantified.	Long lists of items difficult to rank. Items open to different interpretations.
e. semantic differentials	Easily completed by respondents. A useful means of discovering people's predominant feelings. 'Socially acceptable' responses less likely.	Difficult to quantify.
f. questionnaires	Easy to administer to large groups at relatively low cost.	Not easy to secure information on sensitive issues and some responses may not be truthful. Effort required to follow-up non-respondents to ensure adequate response rate.
B. personal products • achievement tests (written)	Easy to administer to large groups at relatively low cost. Can use already developed standardised tests.	Time-consuming to develop and validate. Results don't always reflect ability to transfer and apply knowledge and skills.
• samples of work	Often readily available and easy to collect. Provides a good indication of ability to respond to a specific task within a specific context.	May be time-consuming to assess. Results may not reflect ability to transfer and apply knowledge and skills. Results may be biased by external assistance.

	Advantages	Disadvantages
2. Compiled by an independent observer. pp.100-105		
A. written accounts	Can provide an independent account of program complexity enhancing understanding.	Subject to observer bias. May present difficulties in analysing data. Time-consuming. Observers can change the environment. Inter-and intra observer reliability can be difficult to obtain.
B. observation forms a. observation schedules	Provide a means of collecting data focussed on specific program components. Structured observations provide basis for data analysis.	Subject to observer bias and variations in quality of observations. May overlook important program elements.
b. rating scales	See above.	See above.
c. checklists or inventories	See above.	See above.
C. oral responses either singly or in groups		
a. face-to-face interviews	Allow in-depth probing, and sensitive issues can be discussed.	Costly, time-consuming, subject to interviewer bias, may be more difficult to probe sensitive issues depending on the issue.
b. telephone interviews	Allow in-depth probing, and sensitive issues can be discussed. Less costly and less time-consuming than face-to-face interviews. May reduce some elements of interviewer bias.	Costly, time-consuming, subject to interview bias, may be more difficult to probe sensitive issues depending on the issue.
c. group interactions e.g. focus groups, nominal groups.	Allow exploration of a range of views on a particular issue. Minimal participation of interviewer reduces bias.	Costly, time-consuming to administer and analyse the results. Subject to undue influence by individuals within the group.
search conferences	Encourage unrestricted and creative exchange of ideas about programs and future development.	Requires considerable amount of planning and organisation, and can be costly.

	Advantages	Disadvantages
3. Information compiled using media. p.106		
a. audiotape	Provides permanent record useful for subsequent analysis by one or more people, less intrusive than video-tape.	Very time-consuming to analyse, quality can be variable depending on the setting. Can restrict openness of responses.
b. video-tape	Provides permanent 'observation' record useful for subsequent analysis by one or more people.	Costly, requires specialised skills, and may be associated with practical difficulties in executing depending on the setting. More intrusive than audio-tape and may cause changes in respondent behaviour.
c. time-lapse and still photography	Provides permanent observation record at different points in time. Useful for subsequent analysis by one or more people.	Less intrusive than video-taping, but may cause changes in respondent behaviour. Requires specialised skills.
4. Through use of unobtrusive techniques. e.g. physical traces, archival records, private records, simple observation, contrived observation. pp.106-107	Avoids bias related to the presence of an observer, and people's awareness that they are being studied.	Difficulties in making observations without people being aware of them. Often records are incomplete, unavailable or difficult to decipher. Retrieval can be time-consuming.
5. From existing records. e.g. public documents, files, existing databases. p.107	Unobtrusive, and may be inexpensive.	Existing documents may be disorganised, incomplete or unavailable. Retrieval for analysis can be time-consuming, and in some instances, costly.

Note: In all these information collection methods, both quantitative and qualitative data can be assembled.

FIGURE 2.3: COLLECTING AND ANALYSING EVALUATIVE INFORMATION (STAGE TWO)

A. SELECTING FROM THE RANGE OF INFORMATION COLLECTION ALTERNATIVES p.117

The selection of information collection techniques will depend on the key evaluation issues, available resources and the time-frame which will ensure that the findings are available when they are needed. For all key evaluation questions, we need to have confidence in our findings. Sometimes the collection of information from more than one source (using more than one method) enables us to feel more sure of the validity of the findings.

B. REVIEWING INFORMATION COLLECTION TECHNIQUES p.117

Once information collection techniques have been specified for each key issue in the evaluation, they need to be reviewed along the following lines:

- Will the data collected give a comprehensive picture of what is being evaluated?
- Does the data management plan make effective use of existing data?
- Will the cost of data collection be justified, given the amount and kind of information it will provide?
- Will the information be reliable?
- Can the data collection be carried out without unduly disrupting the program and taking too much of the time of the program providers?
- Are the data collection procedures legal and ethical?
- Can the data be collected and analysed within the time constraints of the study? (Owen, 1993:57)
- Is a pilot study necessary to address some of the questions identified above? (See separate note below on Pilot Study).

C. TESTING INFORMATION COLLECTION TECHNIQUES VIA A PILOT STUDY pp.117-119

A pilot study is a small scale study designed to test the effectiveness and efficiency of selected data collection techniques and instruments. It means trying out the instruments under conditions similar to those expected in the evaluation, and so it should include a range of people who represent those who will participate in the evaluation. It may, for example, be used to test the design of interview schedules, self-report instruments or surveys designed for widespread distribution. It is often carried out prior to a larger scale study to ensure that resources are not wasted via unanticipated difficulties in implementing the data collection strategy and/or the collection of unnecessary or invalid data.

Typical questions addressed by pilot studies are:

- Will the instruments provide the needed information? Are certain words or questions redundant or misleading? Are the instruments appropriate for the people you will be testing?
- Will information collectors be able to use the instruments properly? Can they administer, collect, and report information using the written directions and special coding forms?
- Are the procedures standardized? Is everyone collecting information the same way?
- How consistent is the information obtained by the instruments? (Is it reliable?) (Owen, 1993)
- How accurate is the information obtained with the instruments? (Is it valid?)' (Kosecoff & Fink, 1982:184).

D. FINDING A SAMPLE pp.120-123

'One way of finding out about a group of people is to collect information from everyone in the group. For large groups of people this is prohibitively expensive and impractical. The alternative is to collect information from only some people in the group in such a way that their responses and characteristics reflect those of the group from which they are drawn. This procedure is much cheaper, faster and easier than surveying all members of a group. This is the principle of sampling.' (de Vaus, 1990:60)

There are two basic types of sampling - probability and non-probability sampling:

Types of Probability Sampling: pp.120-121

- simple random sampling
- systematic sampling
- stratified sampling
- multistage cluster sampling

Types of Non-Probability Sampling: pp.122-123

- purposive sampling
- quota sampling
- availability sampling

E. ANALYSIS OF INFORMATION IN EVALUATION pp.123-126

Information analysis means 'processing the data through the interaction of three processes:

- Data reduction, the process of simplifying and transforming the raw information according to some logical set of procedures or rules.
- Data display, the development of an organised assembly of information which leads to drawing of conclusions about the key evaluation questions.
- Conclusion drawing, making meaning of the data in the broader context of the evaluation issue being examined.' (Owen, 1993:57-58)

a. Evaluation as Proof: Implications for Information Analysis

Traditional evaluation techniques are aimed at proving that a given program intervention makes a difference. These involve setting up as controls, alternative versions of the program or intervention being evaluated. In such evaluations, the same information is collected about two groups:

- the experimental group (i.e. the group involved in the program or intervention to be evaluated).
- the control group (i.e. the group that is as similar as possible to those in the experimental group, but which does not receive the program or intervention to be evaluated).

While desirable, setting up true experimental 'control' designs can often be time-consuming, costly and extremely difficult in practice.

b. Evaluation as Investigation: Implications for Information Analysis

An alternative way of viewing the role of evaluation is to conceptualise it, not as a means of providing proof, but as a means of assisting program developers plan and deliver the best possible program interventions. This moves the evaluator into the role of investigator drawing upon both qualitative and quantitative information in response to questions and issues that arise in the course of evaluation.

F. INFORMATION ANALYSIS TECHNIQUES pp.126-128

- the use of summary statistics to describe a situation and to answer key evaluation questions.
- analysis and display of qualitative data.

G. GOOD PRACTICE IN INFORMATION COLLECTION pp.141-144

Utility Standards:
(i.e. those concerned with the provision of useful information)

- Information Scope and Selection

Feasibility Standards:
(i.e. those concerned with the evaluation being realistic, diplomatic and financially well managed)

- Practical Procedures
- Cost Effectiveness

Propriety Standards:
(i.e. those concerned with legal and ethical issues)

- Rights of Human Subjects

Accuracy Standards:
(i.e. those concerned with the validity and reliability of the knowledge)

- Defensible Information Sources
- Valid Measurement
- Reliable Measurement
- Systematic Data Control
- Analysis of Quantitative Information
- Analysis of Qualitative Information

FIGURE 2.4: REPORTING (STAGE THREE)

A. DETERMINE AUDIENCE NEEDS

Identify:

- key concerns
- personal characteristics and preference
- the time-frame for interim and final decisions
- interim and final reporting deadlines

DEVELOP A REPORTING PLAN TO MEET THOSE NEEDS

B. REPORTING OPTIONS pp.133-137

- written versus oral (with or without visual aids).
- progress versus final.
- substantive (main report) versus secondary (such as technical details of data management).
- summary versus main report.
- formal versus informal.
- descriptive versus recommendatory.

POSSIBLE COMMUNICATION FORMS FOR DIFFERENT AUDIENCES

POSSIBLE COMMUNICATION FORM	Report	Executive Summary	Professional Paper	Popular Articles	News Release, Press Conference	Public Meeting	Media Appearance	Staff Workshop	Brochure, Poster	Memo	Personal Discussions
Funding agencies	■	■									■
Program administrators	■	■	■	■	■			■		■	■
Board members, trustees, other management staff		■									
Advisory Committee	■	■	■								
Political bodies (city councils, legislatures)		■		■							
Community groups				■		■					
Current clients				■		■	■				
Potential clients											
Program service providers		■		■				■	■	■	■
Organizations interested in program content			■	■							
Media					■	■					

SOURCE: Adapted from Anderson & Ball (1980:102-3) by Morris et.al. (1987:22). Adapted further by the authors (1993).

C. GOOD PRACTICE IN EVALUATION REPORTING pp.141-144

Utility Standards:

(i.e. those concerned with the provision of useful information)

- Value Interpretation
- Report Clarity
- Report Dissemination
- Report Timelines

Propriety Standards:

(i.e. those concerned with legal and ethical issues)

- Full and Frank Disclosure
- Balanced Reporting
- Fiscal Reporting

Accuracy Standards:

(i.e. those concerned with the validity and reliability of the knowledge)

- Justified Conclusions
- Objective Reporting

FIGURE 2.5: AN EVALUATION PLANNER

Planning: Stage One

1. Purpose of the Evaluation

   Why is the evaluation being done?

2. Evaluation Audience(s)

   Who will review and use the information?

3. Evaluation Resources

   What personpower and materials are available?

4. Evaluation Focus(es)

   Context, design, delivery, outcomes?

Collecting and Analysing: Stage Two

5. Key Evaluation Questions and Data Management

   How can we collect and analyse data around the key issues?

(i) Key questions

For each question, outline the data management techniques to be used.

Question 1	Question 2	Question 3
|	|	|
Data Management 1	Data Management 2	Data Management 3

(ii) Data Management

For each question devise a detailed data collection and analysis scheme.

Reporting: Stage Three

6. Reporting the information from the evaluation.

   What reporting forms are appropriate at what times?

7. Standards of conduct

   Which of the Standards of Evaluation need special consideration in this instance?

8. Budget and timeline

   What will the evaluation cost, and what will be achieved at key time points during the evaluation?

(Adapted from Owen, 1993:53)

Chapter 3 EVALUATION FOR PROGRAM DEVELOPMENT

As indicated in Figure 2.1, evaluation for the purpose of program development is concerned with making decisions about how best to develop a program to meet desired objectives or needs. The program may only be at the 'proposal' stage or alternatively, there may be real concerns about the effectiveness of an existing program indicating the possible need for substantial modification.

Figure 3.1 illustrates the kinds of questions that might be asked by planners at different levels within ALBE operations when tackling this form of evaluation.

In order to address these questions, it is necessary to collect information from a range of different sources using a variety of data collection strategies. One source of information is related research already in existence (research synthesis), another is data that has been previously collected and organised in a coherent form for some other purpose (secondary analysis). Yet another strategy is to locate examples of good practice in the field and to learn through a study of their policies and programs. These approaches can often be usefully combined and/or used as an integral part of 'needs assessment' studies which are perhaps the most common approach to evaluation for program development.

Evaluation for Program Development

A knowledge of needs and available resources is fundamental to making rational decisions with respect to the planning and provision of services. Such knowledge enables planners at all levels to:

- determine priorities for geographic areas, client groups and service delivery
- train and allocate staff appropriately
- locate services and facilities to maximise program effectiveness and efficiency
- substantiate applications for funding and additional resources.

The following five key questions will be useful for the initial development of a needs assessment strategy (adapted from the work of Donaldson et. al. 1987):

1. What are the main objectives in carrying out a needs assessment?

   A needs assessment within ALBE is likely to be directed towards:
   - identifying discrepancies between the desired or ideal level of service provision and the actual level of provision
   - providing reasons for the discrepancies or unmet needs
   - identifying what needs should be given priority for action.

   However, a needs assessment can serve purposes other than these planning and administrative ones. Depending on the methods used, it can lead to an improved understanding of the community being served by the program and may also result in closer links between agencies, both voluntary and statutory, involved in service provision.

2. Who are the key audiences?

   The key audiences for a needs assessment depends on the level at which the needs assessment is done (see Figure 3.1). National and state policy makers, regional co-ordinators, program managers and practitioners may well have different perceptions of what constitutes evidence of need.

3. Whose needs are to be assessed and how are they to be defined?

   Prior to the design and conduct of the needs assessment it is often important to explore the nature and context of the needs in question. Donaldson et.al (1987: 36) suggests that carrying out the following three tasks may be useful in this preliminary work:

- examining the population requiring attention and identifying those groups likely to be the primary beneficiaries of the program proposed. It is often important to ascertain their social, economic, political and demographic characteristics.
- describing the nature of their need in terms of type (economic, psychological, organisational), the scope (how many people are affected), and degree (how severe is the need).
- detailing the background to their need in terms of how it originated and its pattern of development over time.

The evidence of need will depend upon whose responses are sought and the methods used to collect them. As discussed in more detail below, the definition of need can vary and this can have a significant impact on the findings.

4. What questions should be asked?

Frequently, needs assessments attempt to research sensitive issues. The development of creative questions is critical to the accurate detection of how people feel about an issue.

5. What resources are available?

The resources here refer to time, staffing, funding and expertise.

A Closer Look at the Definition of Need

'A need is a value judgement that some group has a problem that can be solved' (McKillop, 1987: 16). Values are central in both the definition and evaluation of need. Researchers and planners have identifed a number of alternative concepts of needs and each can be appropriately used depending on the context of the study. These include normative, comparative, felt and expressed needs (Bradshaw, 1972):

- normative needs are those which are said to exist if a group of people are experiencing something less than a desirable standard or level of service. Within ALBE operations, this standard might be established by policy makers who, for example, might decree that a certain scale of provision was appropriate for a region with a given migrant population.
- comparative needs are not based on a standard but on the relative conditions of one group measured against the conditions experienced by another group. For example, funding bodies may wish to ensure that monies allocated to women's community groups for literacy and basic education provision is equitable to funding provided for college-based mixed groups.
- felt needs are those perceived by groups experiencing the problem and are often referred to as 'wants'.
- expressed needs are wants that are acted upon in the form of 'demand'.

In any one study, heavy reliance on a single definition of need can threaten the validity of the needs analysis as illustrated by the examples below:

- Felt needs are those expectations which the target population hold about their own outcomes (McKillop, 1987:12) and depend heavily on their insight into their own problems. If the purpose of the study was to determine the literacy and basic education needs of unskilled employees in a particular industry, it may well be that the employees themselves are unable to identify the full range of skills or envisage the type of service provision that would enable them to achieve appropriate literacy for their particular context. A lack of expertise in the target population may result in important needs remaining unidentified, needs that would be identified by experts under the normative definition of need.

- A need that is expressed in terms of use of service (demand) does not of itself justify the servicing of that need. Use does not automatically mean that servicing the need is appropriate or should be made a priority. Furthermore, such a definition of need restricts the focus to services that are currently available with the tendency to 'produce solutions that continue the status quo' (McKillop, 1987:12).

Hence, in any one needs assessment, a range of information collection strategies embodying different need definitions might have to be used to build up a composite picture of community needs. This in turn means that decision-making will involve the challenge of integrating multiple sources of information.

In addition to secondary data analysis, a review of related literature and researching good practice in other programs, new information can be sought from key groups which, depending on the needs in focus and the resources available, might be the general population, specified groups within the population (or target population), individuals in direct contact with those experiencing the problems (key informants), service providers or any combination of these.

This information may be obtained via any one or more of the information collection or data management alternatives identified in Figure 2.2: Information Collection Alternatives. These typically include written survey questionnaires, face-to-face interviews with individuals, search conferences, nominal group forums and focus group interviews. Figure 2.2 refers you to a more detailed description of each of these and other information collection alternatives.

Having identified the needs through information collection, analysis and evaluation will require that the value decisions involved in the study are made explicit (McKillop, 1987:16).

Evaluation for Program Development Within ALBE: A Case Study

Figure 3.2 provides a broad framework for a hypothetical study which might be carried out within ALBE. It is used to demonstrate how those at different levels within the field might become collectively involved in program development evaluation. While this is a hypothetical case, the content was drawn from descriptions of actual state level projects funded by the Adult Literacy Action Campaign of the National Policy on Languages (1989). The background to the case is described below:

> Adult Education in South Australia was preparing to embark upon a Workplace Basic Education Project - a scheme whereby classes in reading, writing, oral communication and basic mathematics would be provided for workers during work time in the workplace. Important aims were to develop a workplace basic education philosophy and practice that would be responsive to the changing needs of industry and society and to design a literacy and numeracy program to improve the basic education of working people. Hence, it was important for program planners to locate those workplaces where the need for such a program was greatest, and to tailor a literacy and basic

> education program which would be effective in the workplace context. Due to limited resources, it was agreed that a needs assessment study would focus on two regions only with the possibility that the initiative would be extended to other regions if it was believed that the program should be developed and was proving to be successful in practice.
>
> As part of this initiative, it was decided that a better understanding of communication requirements in the workplace was necessary. This was largely a response to the fact that service providers had begun to realise that an increasing number of people were giving job related reasons for enrolling in adult literacy programs - an indicator they believed reflected workplace communication problems. A group of service providers agreed to appoint a team to identify and research workplace communication with a view to providing input to workplace program design. The resulting project was an in-depth study on communication using a suburban supermarket as the sample workplace. Communication in a social context was observed, with a specific focus on use of language, communication networks and contextual support.

While there are many possible routes to researching the needs implicit in this instance, Figure 3.2 provides an example of how state, regional and provider bodies might focus on different, but interdependent, facets of the evaluative task.

The Impact of Program Development Evaluation

The information provides the basis for setting or refining objectives and allocating resources accordingly. By comparing needs with existing service provision, the disparity between the two can be identified and the relative importance of those gaps in terms of existing objectives and priorities assessed.

In some instances, local providers might respond directly to close the gaps via internal reallocation of staff and resources. Alternatively, requests for additional funding might be justified on the basis of 'evidence of need'. The needs assessment may simply confirm the existing objectives and priorities of the service provider and strengthen applications for continued funding.

Regional bodies might identify high priority areas that are currently not being serviced and reallocate resources accordingly.

State authorities might respond to a needs analysis by providing additional resources for curriculum and professional development in specific areas to substantially alter program delivery.

Based upon a comprehensive study of national needs, both present and future, national authorities may decide that a revision of state organisational structures is a necessary prerequisite for effective service provision and efficient use of resources.

References

Australian Advisory Council on Languages and Multicultural Education (ACLAME), (1989) Adult Literacy Action Campaign: Projects Around Australia, Occasional Paper, Number 3.

Bradshaw, J. (1972) The Concept of Social Need, New Society 30: 610-43.

Donaldson, D.E., Meiller, L.R., Miller, L.C., & Summers, G.F (ed.) (1987) Needs Assessment Theory and Methods, Iowa State University Press, Iowa.

McKillop, J. (1989) Need Analysis: Tools for the Human Services and Education, Sage, Newbury Park, California.

Further Reading

Glampson, A., Scott, T. & Thomas, D.N., (1975) A Guide To The Assessment of Community Needs and Resources, National Institute for Social Work, London, W.C.I.

Neuber, K.A., Atkins, W.T., Jacobson, J.A., & Reuterman, N.A. (1980) Needs Assessment: A Model for Community Planning, Sage, London.

Nickens, J.M., Purga, A.J., & Noriega, P.P. (1980) Research methods for needs assessment, University of Press of America Washington, D.C.

Office of Local Government, Department of Local Government and Administrative Services (1987) Community Development Needs Study of Perth's South East Region, Australian Government Publishing Service, Canberra.

Stufflebeam, D.L., McCormick, C.H., Brinkerhoff, R.O., & Nelson, C.O. (1985) Conducting Educational Needs Assessments, Boston: Kluwer-Nijhoff.

274

275

FIGURE 3.1: EVALUATION FOR PROGRAM DEVELOPMENT - QUESTIONS TO ASK

	QUESTIONS TO ASK	FOCUS OF EVALUATION	KEY AUDIENCES FOR EVALUATION
NATIONAL	What are the current national and Commonwealth agendas, priorities, contexts and goals? What part can ALBE play in achieving national goals? Why do we need a program and what should be its goals? What are the current literacy skills of the population? Are there specific areas of need? Which should receive priority? How should resources be allocated to service the total range of needs identified?	Current situation and contexts, policy and goals. Existing infrastructure, skill levels, present and future needs.	Federal Ministers Portfolio/departments.
STATE/TERRITORY	What are the current priorities, contexts and goals nationally, and in this State/Territory? What part can ALBE play in achieving our goals? What are the particular characteristics of need in this State/Territory? How will the need be met? What will be our policy objectives and priority areas? Is our policy and are our objectives consistent with national policy and other policies in this State/Territory? What types of programs and strategies (e.g. curriculum and professional development) are needed to achieve policy goals and objectives? What infrastructure or changes in infrastructure are required? How/when will these changes be made? What are the key agencies to be involved in these programs or strategies. How will they be involved?	Current situation and context, policy and goals. Existing program provision, infrastructure, present and future needs.	Policymakers, governments, ministers.
REGIONAL	What are the particular needs of this region? What are the region's policies, objectives and priorities? Do they address these needs and are they consistent with state and national policies, objectives and priorities? What is already happening to meet the region's needs? What else would make the difference? Where should we see new provision and/or innovative provision?	Current situation, contexts, policies and goals. Existing program provision, infrastructure, present and future needs.	State authorities, Regional planning groups.
PROVIDER	Are our program policies, objectives and priorities consistent with regional, state and national policies, objectives and priorities? Does our policy provide a framework for the program? What is the present and future need for the program? Do we know who and where our potential program users/students are? How does our program relate to other programs in the area. What are the local conditions that impact on our program now and in the future?	The current situation, context, program policy and objectives. Existing program design, present and future needs.	Funding agencies, policymakers, program planners and deliverers.

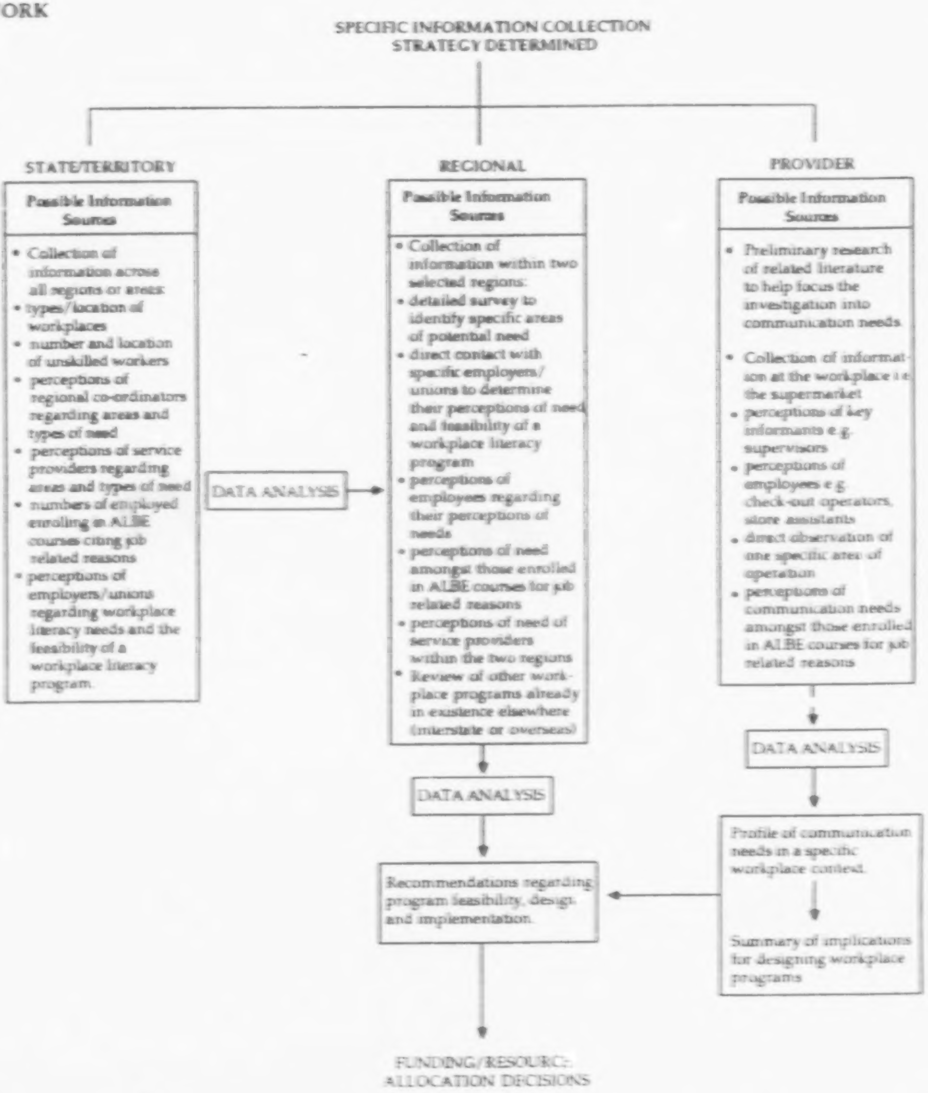
FIGURE 3.2: AN EVALUATION FOR DEVELOPMENT FRAMEWORK

WORKPLACE BASIC EDUCATION PROJECT:
CASE STUDY OF AN EVALUATION FOR DEVELOPMENT

The decision is made to conduct a workplace needs assessment study involving three different levels of operation within ABLE.

PLANNING

	State/Territory	Regional	Provider
Main Objectives	To identify protential areas of adult literacy and basic education need in the workplace across the State.	To identify literacy needs in the workplace in two selected regions. To improve the links between regional planners, service providers and industry.	To identify specific communication needs related to the workplace context.
		To design an effective workplace literacy and basic education program to service the two selected regions.	To provide program guidelines with respect to developing workplace communication skills.
Key Audiences	State Planners Regional Planners	Regional Planners Service Providers	Regional Planners Service Providers
Scope	Preliminary exploratory needs study designed to select two regions likely to have greatest need for the proposed workplace program.	In-depth needs assessment of adult literacy and basic education needs in the workplace within the two selected regions.	Highly focussed study of communication needs within the context of a specific workpace (i.e. a supermarket).



STATE/TERRITORY

Possible Information Sources

- Collection of information across all regions or areas:
- types/location of workplaces
- number and location of unskilled workers
- perceptions of regional co-ordinators regarding areas and types of need
- perceptions of service providers regarding areas and types of need
- numbers of employed enrolling in ALBE courses citing job related reasons
- perceptions of employers/unions regarding workplace literacy needs and the feasibility of a workplace literacy program.

REGIONAL

Possible Information Sources

- Collection of information within two selected regions:
- detailed survey to identify specific areas of potential need
- direct contact with specific employers/unions to determine their perceptions of need and feasibility of a workplace literacy program
- perceptions of employees regarding their perceptions of needs
- perceptions of need amongst those enrolled in ALBE courses for job related reasons
- perceptions of need of service providers within the two regions
- Review of other workplace programs already in existence elsewhere (interstate or overseas)

PROVIDER

Possible Information Sources

- Preliminary research of related literature to help focus the investigation into communication needs.
- Collection of information at the workplace i.e. the supermarket
- perceptions of key informants e.g. supervisors
- perceptions of employees e.g. check-out operators, store assistants
- direct observation of one specific area of operation
- perceptions of communication needs amongst those enrolled in ALBE courses for job related reasons

Profile of communication needs in a specific workplace context.

Summary of implications for designing workplace programs

Recommendations regarding program feasibility, design and implementation.

FUNDING/RESOURCE ALLOCATION DECISIONS

BEST COPY AVAILABLE

Chapter 4 EVALUATION FOR PROGRAM CLARIFICATION

This form of evaluation is concerned with clarifying the underlying logic of a program. 'Logic' here is taken to mean 'a conceptual view of how a given program operates' (Owen 1993:150). Evaluation for program clarification results in a systematic framework or description of a program. It involves describing how a program is supposed to work, identifying cause-and-effect relationships, and in doing so, provides a rationale for the program as an intervention (Smith, 1989:16). Examples of program clarification questions that might be asked at different levels within ALBE operations are listed in Figure 4.1.

Clarifying the logic of a program can provide the basis for:

- planning a new program
- increasing commitment to an existing program amongst staff who, in the process, develop a much clearer understanding of intended outcomes and the strategies to achieve them
- improving the quality of decisions about subsequent program modification because these decisions are based on an explicit identification of program strengths and weaknesses
- setting up a valid monitoring evaluation process for the program in action
- planning a subsequent impact evaluation.

Drawing Up A Program Logic Statement: Two Models

The level of detail and the format of a program logic statement may vary significantly depending on the program in focus and the purpose of analysis. Two approaches to developing a program logic statement have been included here for the following reasons:

(i) the use of 'logical operators' (for example, AND, IF, THEN, OR, etc.) to describe the essential features of a program and links between them. Such a logic statement is relatively easy to conceptualise and construct, especially by those who have first hand knowledge of the program in operation.

(ii) the use of a hierarchy of objectives model, in which a given objective is dependent on a prior one. This highlights those objectives and related program components which may or may not be amenable to subsequent impact evaluation.

(i) The Logical Operator Model

Figure 4.2 provides an example of an 'And/If/Then' program logic statement for a proposed workplace basic education program - a scheme whereby classes in reading, writing, oral communication and basic mathematics would be provided for workers during work time in the workplace. From this, it can be seen that logic statements attempt to show how resources and activities are directed towards achieving specific program outcomes and broader program outcomes. The emphasis is on cause and effect relationships; 'if these inputs (people, resources and activities) are combined in this way, then these are the expected outcomes'. The logic statement is thus concerned with the internal coherence of the program. It makes explicit the underlying assumptions implicit in program design and might be regarded loosely as a 'logic of action'.

The advantage of conceptualising the logic of a program in this kind of descriptive framework is that it highlights those activities which must take place before there is any reason to believe that the intended impact (processes and outcomes) would occur. In the case of a proposed program, conceptual gaps can be rectified via logic, discussion and policy analysis. In the case of a program already operating, information gaps can be addressed via an investigation (Patton, 1986: 57-58) and refinements can be made to policies, objectives and implementation strategies.

A Program Clarification Framework

One approach to program clarification, outlined below, involves cooperative planning. This requires stakeholders to consider program design issues and participate in the documentation of the existing program. An integral part of the process is the development of a program logic statement.

Figure 4.3 provides guidelines for developing a program logic statement as an integral part of program clarification evaluation. The major steps identified in the framework are discussed here:

1 Someone with evaluation expertise acts as leader of the evaluation. The leader needs to negotiate an **agreement** from stakeholders about the purpose of the evaluation. A small team selected from program personnel, managers and deliverers, works with the leader on the evaluation. Stakeholders and program personnel should be persuaded of the benefits of the evaluation.

2. The **boundaries** of the program in focus need to be defined. There is no one definition of boundaries. Geographical area, goals, services, or benefits provided, clients served, or a combination of these might be used (Smith,1989:39). Regardless, because of the time and resources which must be devoted to program clarification, the program in focus should be a 'major program' i.e. a program which has a substantial amount of resources directed towards a problem or need (Smith, 1989: 42), and the decisions which are to be affected should be important ones.

3. All **documents** related to the program should be analysed. These may include legislation, hearings, debates, reports, memoranda, funding applications, program and resource statements. It is important to remember that program statements in documents may vary depending on their audience and that any inconsistencies or conflicting statements may reflect this, and/or confusion among program managers and deliverers.

4. Based on the documentation, a **program logic statement** is constructed. As indicated earlier, there are a number of approaches for displaying a program's logic but regardless of the approach selected, the basic statement shows the essential components of the program connected in a causal sequence to ultimate outcomes. Different parts of the program should be readily identified to facilitate analysis and presented in a form which is easily understood by stakeholders (Smith, 1989: 53). Ideally, construction of the statement should involve key program staff who have a close working knowledge of the program in practice and who stand to benefit most from developing the logic. Assessment of program plausibility is carried out at a later stage (see Step 9).

5. Having prepared the program logic statement, the next step is gather **stakeholder perceptions of the program**. Stakeholders are 'those persons or groups who impact a program in very significant ways or who are similarly affected by the actions of a program' (Smith, 1989:82). Therefore, stakeholders might include funding and legislative bodies, boards of study, management committees, tripartite workplace committees, public interest groups, program managers, program staff and their clients. Program information from stakeholders is usually collected via personal interviews which focus on their:
   - perceptions of an existing program
   - assessments of the need for and relevance of stated goals of an existing or proposed program

- concerns about how an existing program is being implemented and whether it is accomplishing its goals
- assessments of the adequacy of resources for a planned or existing program
- opinions regarding needs for further evaluative information on the planned or existing program (Smith, 1989: 84).

(See Chapters 7 and 8 for a discussion of interview data collection techniques and approaches to interview data analysis.)

6-8 Stakeholder perceptions of the program are then compared with the program logic statement. Any differences are identified, validated with additional data where necessary and **refinements made to the program logic statement**.

9. Once this is achieved, input from stakeholders is sought regarding the **plausibility of program design**. Plausibility is defined as the existence of necessary and sufficient conditions for a program to succeed. These include a clear intention to bring about outcomes, appropriate activities and adequate resources (Owen, 1993:153). Figure 4.4 provides a series of questions for reviewing program plausibility.

10. Once program plausibility has been assessed, the findings need to be drawn together into a series of **conclusions and recommendations**.

(See Chapter 9 for further discussion on making recommendations when reporting on evaluation findings.)

11. While it should be emphasised that program design clarification is often an important activity in its own right, the findings may prompt a range of **decisions** about the program. These may lead to:

- withdrawal of or further support for a proposed program
- changes in an existing program
- a monitoring or impact evaluation, if the program as implemented is consistent with the refined program plan.

(ii) An Expanded Objectives Hierarchy Model

As an alternative to the Logical Operator approach to clarifying the logic of a program, objectives can be used as the focus. In a hierarchy of objectives, it is assumed that the achievement of each lower objective is necessary before the next higher objective could be expected to be achieved. 'In this format the pattern of relationships between the results to be achieved, and thus cause and effect reasoning behind the program, can be made explicit, providing a description of the program's essential values and assumptions.' (Milne, 1993:42)

Figure 4.5 shows a simple objectives hierarchy designed to convey the logic of the proposed workplace basic education program which was depicted earlier in the Logical Operator model (Figure 4.2). The objectives are classified into three levels: immediate, intermediate and ultimate. Alongside each of the objectives, an explanation of the assumptions underlying them is provided.

It is possible to expand this model further by identifying those program strategies and activities which are directed towards achieving each of the objectives, and then listing the types of performance information required to signal the extent to which the program has succeeded in achieving them. Figure 4.6 gives some examples of how this might be done for both the immediate goals and the intermediate goals of the proposed workplace basic education program. The omission of ultimate objectives from this table is intentional - it reflects the enormous difficulty in drawing direct linkages between specific program activities and ultimate goals, and in collecting valid information about, for example, the impact of workplace literacy programs on national productivity levels (Ultimate Objective No.2), or the extent to which participants go on to lead fulfilling lives as a result of their increased skill levels (Ultimate Objective No.1).

Linking Program Logic Statements With Program Monitoring

Like the Logical Operator model, the expanded objectives hierarchy model conveys the underlying logic of the program and both can lead to program monitoring.

For example, planning a program using the expanded objectives hierarchy model has considerable merit. It requires program managers (in consultation with stakeholders) to think about the results they want to achieve in a cause-effect chain and to consider what they should be measuring as the program develops over time. Identifying the performance information required for each objective provides the basis for considering which objectives (and which program components) are more amenable to monitoring than others. Once decisions about performance information are made, reporting requirements can be structured accordingly.

With the growing reliance on strategic management and the demand for rapid responsiveness within the public sector generally, this kind of 'conceptual use' of evaluation for communicating, planning and monitoring is likely to become increasingly important.

(Refer to Chapter 6 for a more detailed discussion on program monitoring.)

References

Milne, C. (1993) 'Outcomes Hierarchies and Program Logic as conceptual tools: five case studies', Paper presented to the Australasian Evaluation Conference, Brisbane.

Owen, J.M. (1993) Program Evaluation Forms and Approaches, Allen and Unwin, St. Leonards, NSW.

Patton, M.Q. (1986) Utilization-Focused Evaluation, 2nd Ed., Sage Publications, California.

Smith, M.F. (1989) Evaluability Assessment A Practical Approach, Kluwer Academic Publishers, Massachusetts.

Workbase NZ Outline for Literacy Audit, A Project of the Adult Reading and Learning Federation, Auckland, New Zealand.

Further Reading

Bickman, L. (1987) Using Program Theory in Evaluation, San Francisco, Jossey-Bass.

Suchman, E.A. (1967) Evaluation Research: Principles and Practice In Public Service And Social Action Programs, New York, Russell Sage.

Wholey, J.S. (1979) Evaluation: Promise and Performance, Washington D.S., The Urban Institute.

Wholey, J.S. (1987) 'Evaluability assessment: Developing program theory. In L. Bickman (ed.) Using Program Theory In Evaluation, New Directions for Program Evaluation, No. 33, San Francisco, Jossey-Bass.

FIGURE 4.1: PROGRAM CLARIFICATION - QUESTIONS TO ASK

	QUESTIONS TO ASK	FOCUS OF EVALUATION	KEY AUDIENCES FOR EVALUATION
NATIONAL	What are the objectives of our policies? Do our policies have internal coherence? Do our policies provide support for program implementation by states and territories? Are our policies plausible in terms of guidance for practice?	Plan	Federal Ministers Portfolio/departments. State/Territory policymakers
	Does the national infrastructure support program implementation in line with policies? Do resource allocation practices reflect our stated objectives and priorities? Does our manpower planning adequately support program development?	Existing Program	
STATE/TERRITORY	What are the objectives of our policies? Do our policies have internal coherence? Do our policies provide support for program implementation? Are our policies plausible in terms of guidance for practice?	Plan	State/Territory planners Funding bodies
	Does the existing infrastructure support program objectives? Will it be adequate to support proposed program changes? What resources are needed to achieve revised program objectives? Do communication and consultative processes support the achievement of program objectives? Are there strategies in place to gain community support for the program?	Existing Program	
REGIONAL	What are the objectives of our policies? Do our policies have internal coherence? Are our policies plausible in terms of guidance for practice?	Plan	Regional staff Regional advisory groups Providers
	Does the existing infrastructure support program implementation consistent with program objectives? Is resource allocation amongst providers consistent with policies and priorities? Do program changes need to be made to achieve current objectives? Are the objectives of staff development programs consistent with curriculum objectives at provider level?	Existing Program	
PROVIDER	What are our objectives and are they consistent with regional objectives? Are they realistic and how are our programs structured to achieve these objectives?	Program Design	Program planners Program staff and clients
	Do our programs offer sufficient contact time to enable students to gain the desired level of skills? Is the present balance of reading and mathematical work appropriate? Which of the present programs should be targeted for an impact evaluation?	Processes	

FIGURE 4.2: A PROGRAM LOGIC STATEMENT FOR A PROPOSED WORKPLACE BASIC EDUCATION PROGRAM

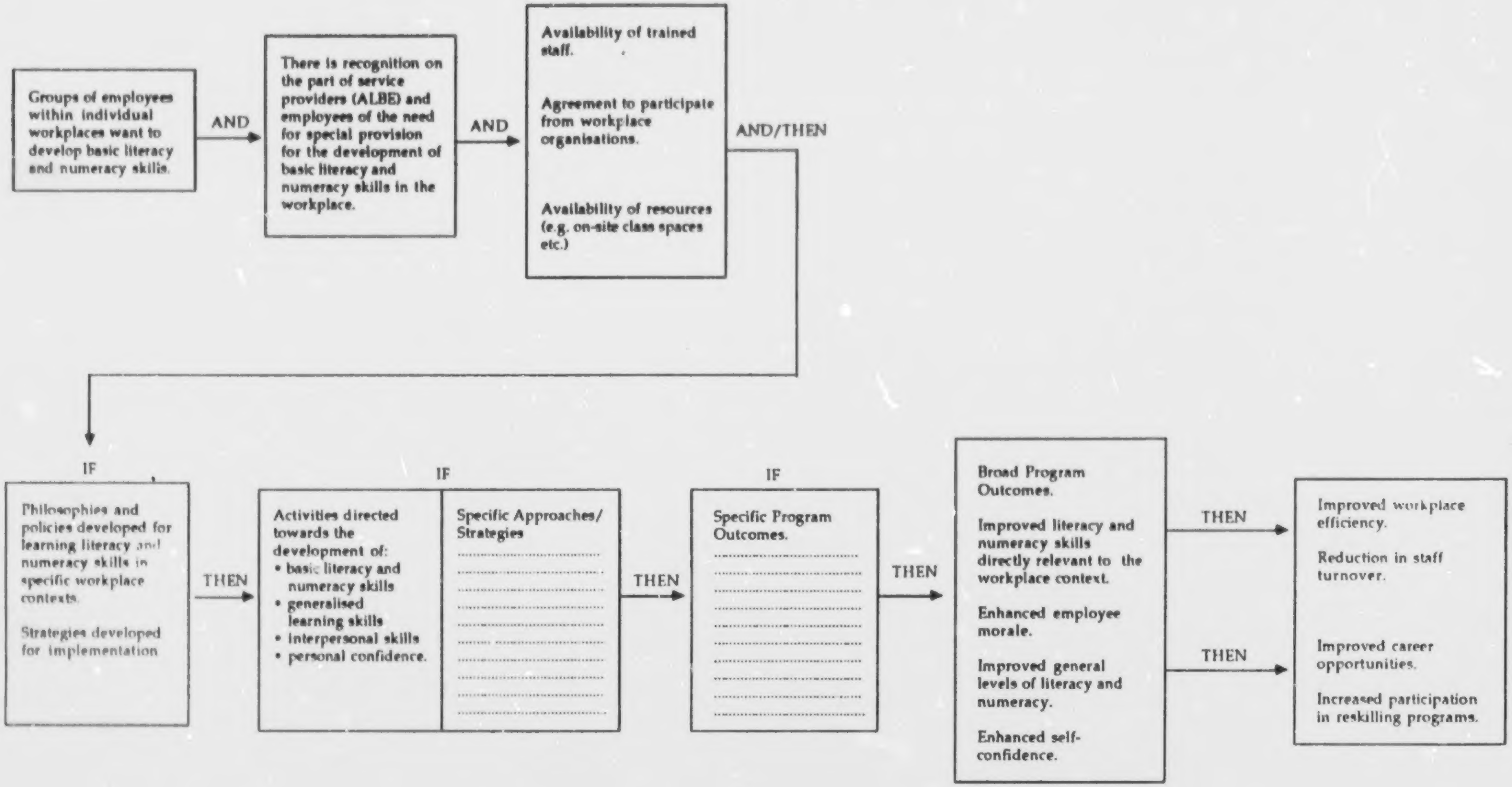

Groups of employees within individual workplaces want to develop basic literacy and numeracy skills.
AND
There is recognition on the part of service providers (ALBE) and employees of the need for special provision for the development of basic literacy and numeracy skills in the workplace.
AND
Availability of trained staff.
Agreement to participate from workplace organisations.
Availability of resources (e.g. on-site class spaces etc.)
AND/THEN
IF
Philosophies and policies developed for learning literacy and numeracy skills in specific workplace contexts.
Strategies developed for implementation
THEN
IF
Activities directed towards the development of:
• basic literacy and numeracy skills
• generalised learning skills
• interpersonal skills
• personal confidence.
Specific Approaches/ Strategies
THEN
IF
Specific Program Outcomes.
THEN
Broad Program Outcomes.
Improved literacy and numeracy skills directly relevant to the workplace context.
Enhanced employee morale.
Improved general levels of literacy and numeracy.
Enhanced self-confidence.
THEN
THEN
Improved workplace efficiency.
Reduction in staff turnover.
Improved career opportunities.
Increased participation in reskilling programs.

FIGURE 4.3: A PROGRAM CLARIFICATION FRAMEWORK

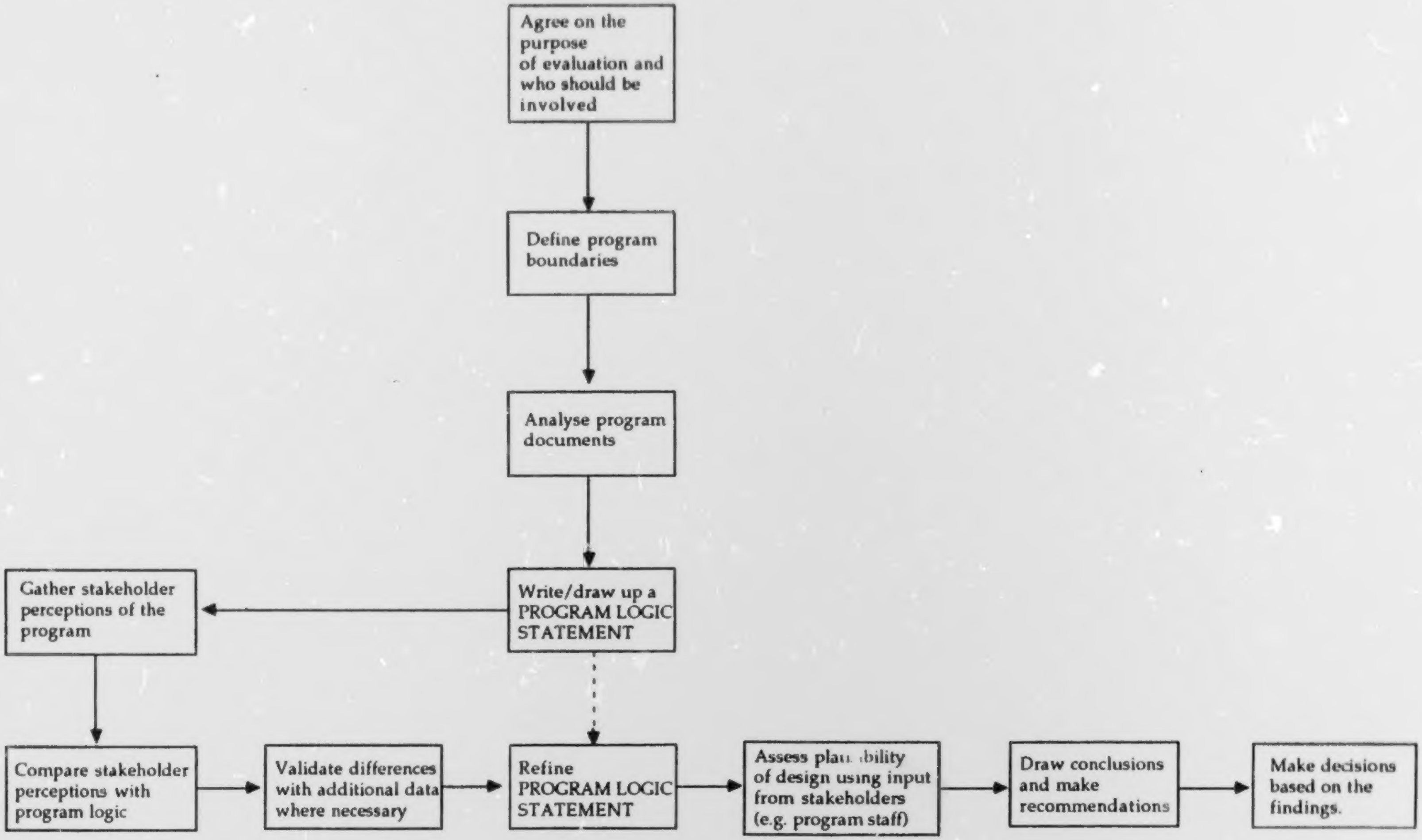

Agree on the purpose of evaluation and who should be involved
Define program boundaries
Analyse program documents
Write/draw up a PROGRAM LOGIC STATEMENT
Gather stakeholder perceptions of the program
Compare stakeholder perceptions with program logic
Validate differences with additional data where necessary
Refine PROGRAM LOGIC STATEMENT
Assess plausibility of design using input from stakeholders (e.g. program staff)
Draw conclusions and make recommendations
Make decisions based on the findings.

FIGURE 4.4: REVIEWING PROGRAM PLAUSIBILITY

Are Objectives Well Defined?

1. Are objectives clearly stated?
2. Have indicators of successful performance (criteria and standards) been specified?
3. Have the sources of evidence for performance been identified?

Are Program Components/Activities Well defined?

1. What is the known purpose?
2. What are the supporting activities?
3. What resources are there to support these activities?
4. Are there identifiable indicators of successful performance (criteria and standards), and
5. Have the sources of evidence for performance been identified?

Are Program Components/Activities Sufficient?

1. Is the content appropriate?
2. Is there sufficient information for the client to learn or do what is expected?
3. Can the client acquire sufficient skill/knowledge in the time-frame to effectively apply it later on?
4. What evidence is there that clients have the necessary pre-requisite skills/knowledge when they enter the program?
5. What agencies/groups are important to the success of the components/activities and what part do they need to play?
6. Are strategies in place to successfully justify the worth of the components/activities?
7. Are promotional efforts pertinent to the targeted audience?

Are Resources Adequate?

1. Are the type and amount of resources clearly identified?
2. Are the resources available?
3. Is there confidence that the necessary resources will be available when needed?

Adapted from Smith, M.F. 'Evaluability Assessment', Kluwer Academic Publishers, 1989: 116-118.

FIGURE 4.5: AN OBJECTIVES HIERARCHY FOR A PROPOSED WORKPLACE BASIC EDUCATION PROGRAM

Hierarchy of Objectives	Underlying Assumptions
III. Ultimate Objectives	
9. To meet the literacy needs of adults - needs which relate both to their work and other life experiences.	Adults whose literacy needs are met will lead fuller, more satisfying lives.
8. To maintain and improve national productivity by raising literacy skills in the workforce.	The modern industrial setting with its rapidly changing technology demands a flexible and educated workforce. The acquisition of literacy skills is essential for workplace efficiency and the re-training of employees.
7. To facilitate and legitimise the establishment and maintenance of workplace literacy programs.	Planning bodies and employees will support workplace literacy programs once their benefits have been demonstrated.
II. Intermediate Objectives	
6. To enhance employee self-confidence and interpersonal skills.	Self-confidence and interpersonal skills are an integral part of the effective use of literacy skills in the workplace.
5. To improve employee morale.	Employees will respond positively to organisations that provide on-site literacy programs which meet their needs.
4. To provide programs that meet the literacy needs of employees in their specific workplace contexts.	Employee literacy skills are most effectively addressed via on-site programs tailored to their needs in the specific workplace context.
I. Immediate Objectives	
3. To train practitioners to undertake literacy programs designed specifically for workplace contexts.	Workplace literacy programs are most effective when they: • involve management, unions and literacy provider in the setting up process • are custom built following a careful literacy audit and training needs analysis • are presented as part of an overall workplace training strategy rather than a remedial service • balance vocational skills development with the skills necessary for broader social interaction • become embedded as an ongoing feature of the workplace training strategy. (Ref: Workbase NZ outline for Literacy Audit)
2. To delineate organisational/individual employee learning needs.	
1. To create an awareness of and support for the creation of on-site literacy acquisition programs amongst employers, unions and potential clients (employees).	

FIGURE 4.6: AN EXPANDED OBJECTIVES HIERARCHY FOR A PROPOSED WORKPLACE BASIC EDUCATION PROGRAM

Hierarchy of Objectives	Strategies/Activities	Performance Information
II. Intermediate Objectives		
6. To enhance employee self-confidence and interpersonal skills.	Conduct individual and group learning activities designed to foster confidence and interpersonal skills.	Observations of changes in employee behaviour made by managers/supervisors
5. To improve employee morale.	Provide positive feedback.	Reduction in staff turnover
4. To provide programs that meet the literacy needs of employees in their specific workplace contexts.	Tailor program content and activities to address the immediate literacy needs of employees.	Assessment of student reading/writing/mathematical skills on work related tasks Course completion rate Attendance rate Student satisfaction rating No. of student hours required for course completion
I. Immediate Objectives		
3. To train practitioners to undertake literacy programs designed specifically for workplace contexts.	Conduct workshop sessions to train tutors to conduct literacy programs designed to meet needs identified in specific workplace contexts.	Level of staff satisfaction Course completion rate Attendance rate
2. To delineate organisational/individual employee learning needs.	Examine company job task breakdown sheets, observe workers at work, record reading/writing tasks undertaken on the job, collect reading/writing material which pertains to the job, interview potential participants to assess personal literacy learning needs, etc.	Student rating of program Employer satisfaction rating Course completion rate Attendance rate
1. To create an awareness of and support for the creation of on-site literacy acquisition programs amongst employers, unions and potential clients (employees).	Discussions with management/unions/employees regarding benefits of on-site workplace literacy programs.	Number of organisations willing to participate Employee enrolment: places offered Enrolments: applicants Gender ratio Penetration rates (e.g. of designated priority groups)

Evaluation for program improvement (or process evaluation) focusses on program delivery. It is concerned with understanding the day-to-day reality of programs in practice and using that understanding to improve the quality of program delivery. In-depth and detailed program descriptions are often required in order to make judgements of quality, and hence qualitative data collection techniques are commonly used. Process evaluation typically includes not only a study of the formal activities, but also 'informal patterns and unanticipated consequences' (Patton, 1987: 24).

Process evaluations address the following kinds of questions:

- What is happening in this program?
  - What activities are staff involved in?
  - What activities are clients (students etc.) involved in?
  - What is the nature of staff-client interaction?
- Is this program being delivered in a way which is consistent with program goals and objectives?
- What are the strengths and weaknesses of program delivery?
- What should be changed to make delivery more effective?

Figure 5.1 lists examples of process evaluation questions that might be asked at different levels within ALBE operations. Process evaluations, with their emphasis on description, have a range of uses depending on their audience. For example, they can be used by:

- planning bodies, when making decisions about which type of services to encourage, or whether to develop similar programs at other sites.

- program managers, when making decisions about which program components need to abandoned, modified or preserved.

- service providers, as a means of informing others about program activities, exchanging new program ideas, and identifying gaps in service provision.

- external groups such as funding bodies and public officials, to enable them to understand how a given program works in the field, and to relate more effectively to programs and program staff.

Basic information about program operations is fundamental to evaluation for improvement. Unless evaluation questions are framed within the context of an understanding of program implementation, there is the real danger that they will be the wrong questions and involve the collection of irrelevant or erroneous information.

Process evaluation is carried out in three different contexts (Owen, 1993: 129):

- when stakeholders need information about the program to check its implementation as part of a review of its impact (**implementation studies**).

- in the early stages of a program when there is likely to be on-going development and change. In this instance, as staff experiment with new methods and adapt to local circumstances, the evaluation can provide feedback on program processes and effects on participants (**refining developing programs**).

- when practitioners have particular concerns about the quality of provision at individual site level. In this instance, evaluation is responsive to the needs of those directly affected (**site-level evaluation**).

Implementation Studies

A major reason given for the ineffectiveness of human service programs to meet the standards expected of them is failure at the implementation, or delivery stage. Typically, much evaluation effort has been focussed on outcomes, and relatively little on the extent to which the service is actually being implemented in the form of an operating program (Patton, 1989: 125).

...'Where outcomes are evaluated without knowledge of implementation, the results seldom provide a direction for action because the decision maker lacks information about what produced the observed outcomes (or lack of outcomes). This is the "black box" approach to evaluation.' (Patton, 1987: 27)

Implementation studies are important because it cannot be automatically assumed that a given program is actually being implemented or that it is being implemented as intended. These studies may focus on either implementation and outcomes, or the extent of implementation only:

- Implementation and outcomes (process-outcome studies). Process-outcome studies are often used to justify the worth of the program for an external audience, but they may also be used for policy making (Owen, 1993:131). If, for example, a unique approach to workplace basic education is being used within a particular region, a process-outcome study of the program may be used in deciding whether or not to adopt the program state-wide.

- Extent of Implementation. The primary concern here is the extent to which the program in action reflects the essential features of the program plan. This is called the fidelity approach. The methods used in the fidelity study might include monitoring essential features via a standard checklist, self-reports and interviews. However, because programs are frequently adapted to local conditions and client needs, these studies may also need to be open-ended and capable of describing ways in which the program was adapted, and reasons for this adaptation. It may be important, for example, for policy makers to understand these reasons and those which contribute to fidelity implementation (Owen, 1993:132-3).

Refining Developing Programs

Process evaluation can be useful when a program is in the prototype stage, before being disseminated for general use. The evaluation focusses on limited trials of program activities. Evaluation in this context is described as 'formative' and needs to be flexible in that programs are likely to undergo a process of adaptation to local conditions and client needs. Formative evaluation involves:

- keeping an accurate description of the program and its development
- monitoring implementation, examining and reporting on progress
- using feedback from the evaluation to make adjustments in the program. (King et.al., 1987:17)

Whether formative evaluation is undertaken by external evaluators, or those within an organisation, the data collection methods are often less rigorous than those used in implementation studies. These may include informal meetings and anecdotal evidence which lead to verbal, as opposed to written, reports. Naturalistic enquiry methods (see discussion below) are often used in the evaluation of developing programs because they allow the focus of the evaluation to shift depending on which issues/questions are found to be important.

Site-Level Evaluation

In this context, the planning of evaluation is done at site or local level, and involves identifying the major stakeholders directly associated with the program and responding to their concerns

(responsive evaluation). It is oriented towards program activities, addressing questions such as:

- How is this service or activity going?
- Is it working ?
- How is it affecting our target or clients?
- What is its value and how can we improve it ? (Owen, 1993: p.137)

Responsive evaluation may begin with an immediate or obvious problem or may examine a series of concerns that stakeholders agree should be investigated. These concerns are identified via direct face-to-face contact with people in and around the program and an examination of program documents. Understanding the dynamics of the program is important - this requires a gathering of stakeholder perceptions and involves detailed descriptions of program content and context. Recognising this, Parlett and Hamilton (1972) put forward an evaluation strategy which they call 'evaluation as illumination'. The strategy involves three stages: observing the educational setting, selecting themes through progressive focussing and intensive inquiry and, finally, analysis and explanation.

Responsive and illuminative evaluation strategies are consistent with **'naturalistic enquiry'**, one of two major approaches to site-level process evaluation. The central features of naturalistic enquiry are direct observation in the program setting and an openness to issues and outcomes that emerge. Data collection strategies cannot be specified in advance and must be adaptive and creative. Similarly, careful consideration of appropriate forms of reporting is required in order to communicate effectively with each of the different stakeholder audiences (see Chapter 9 for a detailed discussion of reporting forms).

Naturalistic enquiry may lead to a re-definition of local program goals and changes in program activities. In doing so, it could also instigate the need for related **action research**, a second major approach to site-level evaluation. Action research has as its basis the search for solutions to on-the-job problems (Owen, 1993:141). For example, close observation of a literacy tutor program might reveal deficiencies in the techniques for imparting curriculum development skills. Solutions to remedy these might be discussed and subsequently, program changes made. It would then be highly desirable to monitor the effectiveness of these program changes and their impact on tutor behaviour in the field via action research.

Action research has been described as a cyclical process consisting of the following components (Wadsworth, 1991):

- Reflection on current action. This involves observing discrepancies between what is actually happening in the program compared to what was intended. It involves such questions as: What is going well or not going well? What would we like to see happening? Why aren't we doing what was planned?

- Design. This requires articulating the problem and the questions that it raises, as well as identifying who owns the problem, who stands to benefit from addressing it (i.e. the critical reference group), and who should be involved.

- Fieldwork. This is the data collection stage which begins with open-ended questions which become more focussed later and is particularly concerned with exploring the perceptions of critical reference group members.

- Analysis and conclusions. This involves developing meaning from the findings and generating new insights or theories about what is being done. Feedback is sought from those involved in the evaluation and critical reference groups in particular, to validate the findings and the resulting conclusions.

- Planning. This involves the use of the findings to consider options for improved practice.

304

305

While naturalistic enquiry and action research are heavily oriented towards program delivery, the findings of evaluation may well lead to changes which impact on program plans and outcomes. 'A key feature of site-level evaluation is to build evaluation into the fabric of decision making over time' (Owen, 1993: 139).

Wadsworth (1991:47-60) discusses the need to develop a 'culture of evaluation' whereby evaluation becomes a way of thinking which permeates every kind and level of daily action. To help achieve this, she suggests a number of opportunities for developing a comprehensive program of in-built evaluation:

- Daily informal personal reflection

  Good practitioners intuitively reflect on their teaching practices and modify them accordingly on a lesson-by-lesson and day-to-day basis.

- Weekly Reviews

  These might involve making special diary notes on the progress of particular students, reviewing the week's diary to reflect on priorities and time management issues, arranging informal discussion time with colleagues to discuss facets of program operations.

- Special effort evaluations of particular aspects of practice.

  These will be evaluations which respond to the immediate concerns of those involved in program activities. For example, there may be some concern that the program is not satisfactorily reaching those who need it most. This might prompt an investigation of current strategies for reaching the target group, and might in the initial stages lead to an evaluation of the program newsletter.

- Monthly collective problem-pooling sessions

  Such sessions could provide a forum for raising new issues, identifying those issues clearly in need of action, or tabling contentious issues for further discussion and/or research.

- Annual 'what-have-we-achieved?' and 'where-are-we-heading next year?' workshops

  This might involve reports from individuals reflecting on the program over the year, brainstorming sessions about future program development and identifying action that needs to be taken.

- Comprehensive program stocktakes every 3-10 years

  Figure 5.2 identifies a range of program components that might be included in a comprehensive program review (Adult Literacy Volunteer Tutor Program Evaluation Kit, 1989).

Integrating Evaluation Into Program Delivery

Two case studies are now presented to demonstrate the utility and cost-effectiveness of integrating evaluation into on-going program delivery. The first case study described by Patton (1988:87-88) provides a simple example of how evaluation can be built into program design, facilitating the collection of data and enhancing program provision. The second case study involves the introduction and implementation of a workplace basic education program and demonstrates how the evaluation design is used to provide a focus for program activities and reporting.

(i) Case Study A: An Adult Training Workshop

In this case study based on a 'training workshop', evaluation is first presented as separate from and independent of program delivery (Scenario 1). The problems with this approach are highlighted and then an alternative scenario is described, one in which evaluation is integrated into program design (Scenario 2).

Scenario 1:

> 'Consider.......a typical training workshop where the participants are to demonstrate changes in knowledge, skills, and attitudes as a result of the workshop. The evaluation design calls for a pretest and a posttest. To implement this design, the pretest is administered as the very first item in the workshop agenda. Participants are assembled and told, "Now, before we begin the workshop, we want to administer a pretest so that we can find out how much you have benefited from the workshop".' (Patton, 1988:87)

This approach immediately introduces measurement problems associated with the effects of pretesting and posttesting. Those participants who take a pretest may perform better in the program than those who don't, because the pretest may increase awareness, stimulate learning, or enhance preparation for program activities. For studies requiring high internal validity, a complex design would be needed to isolate the possible contamination effects of pretesting and posttesting and this would include randomly selected and assigned control and treatment groups. This would be both expensive and unwarranted in most site-level evaluations which are directed towards providing decisionmakers with feedback about the 'reasonable effects' of their programs. Therefore, an alternative scenario which fully integrates evaluation data collection into the program is proposed:

Scenario 2:

> 'In this scenario, the workshop begins as follows: "We are ready to begin the workshop. The first part of the workshop involves your participating in a self-assessment of your knowledge, skills, and attitudes. This will help you prepare for and get into thinking about the things we will be covering today in the workshop".
>
> The workshop then proceeds. At the end of the day, the workshop presenter closes as follows: "Now the final workshop activity is for you to assess what you have learned today. To facilitate that process, we are going to readminister the self-assessment that began the workshop. This will serve as a review of the workshop's content and allow you to identify what you have learned. In addition, we should be contacting you in six months to repeat this assessment, both by way of reminding you of what you learned today and of finding out how well you have retained what we have covered".' (Patton, 1988:87-88)

In this second scenario, the evaluation pretest and posttest become a direct part of program design and no control groups are required. The self-assessment instrument serves the dual function of informing participants about the content of the course thus preparing them to learn, and providing evaluation baseline data. The posttest and the six-month follow up not only serve evaluation requirements but reinforce the learning.

The two scenarios reflect fundamentally different ways of thinking about and conducting evaluation. The advantage of integrating evaluation into program design/delivery is that evaluation data collection becomes more cost-effective and less burdensome for program staff who often see evaluation as a separate task imposed on them.

(ii) Case Study B: A Workplace Basic Education Program

A Workplace Basic Education Program is used here as a case study to illustrate the principle of introducing evaluation into project management from its inception. While this case study is hypothetical, some of the content is drawn from actual evaluation work done in the literacy field. As the introduction of the program is described, the merits of conceptualising evaluation as part of program planning and development are highlighted.

The aim of the project was to introduce and evaluate a Workplace Basic Education program at a number of selected sites in order to:

- develop guidelines for good practice related to design, implementation and evaluation of workplace literacy programs,
- document practical steps required to undertake locally negotiated workplace learning programs.

First, a steering group was formed and discussions took place with management and unions to reach a preliminary assessment of literacy needs, the potential target group and the current training environment.

Then a 'literacy audit' was carried out to get a clear picture of company literacy needs and individual worker learning needs. This involved:

- an examination of company job task breakdown sheets,
- observation of employees at work,
- making records of reading/writing tasks,
- collecting reading/writing material pertaining to the job,
- talking with line managers and union officials to identify the impact of literacy difficulties on the job,
- interviewing potential participants to assess personal literacy learning needs.

The literacy audit was also quite explicitly an intervention in itself in that it fostered support for the program amongst both employees and employers. In addition, this initial survey served the critical evaluation function of establishing baseline data. Subsequent data on the impact of the proposed program on worker skills and their ability to function within the workplace could be compared against this baseline for evaluation purposes. Such data collection was practical and cost-effective because it served program as well as evaluation purposes.

The findings were then reported back to management and unions through an in-house steering group with recommendations for the workplace program. Contracts were drawn up in which courses were described in detail, including aims, length and level of course, number of practitioners, specific curriculum content and methods of assessment. As part of this phase, there were several days of discussion amongst key stakeholders about criteria that could be used to determine the degree to which the program had been successful. These criteria constituted a set of questions and outcomes, including both qualitative and quantitative indicators.

An in-house steering group was charged with an evaluation of the resulting program. The group decided that implementation would be the primary focus and this decision contributed substantially to keeping the data collection from being sidetracked into other areas. A key input was information collected at regular program staff meetings. The feedback became the basis for making on-going program modifications. Thus the evaluation was used to improve program delivery. Focussing the evaluation also streamlined the assembly of relevant data for a series of reports on the introduction of the program for a wider audience.

This case demonstrates the value of considering, at the planning stage, the range of opportunities that might be used to integrate evaluation into on-going program delivery. The advantages of effective integration include:

- a reduction in evaluation costs
- improvements in program design and implementation
- a greater willingness on the part of program staff to collect information that will serve evaluation purposes because the information also serves program needs
- an expanded reservoir of evaluation information that can be used to address accountability requirements.

References

Adult Literacy Volunteer Tutor Program Evaluation Kit (1989) Adult Basic Education, Ministry of Advanced Education and Job Training, and Ministry Responsible for Science and Technology, Provincial Curriculum Publications, Richmond, British Columbia.

King, J.A., Morris, L.L. & Fitzgibbon, C.T. (1987) How to Assess Program Implementation, Sage Publications, Newbury Park, California.

Owen, J.M. (1993) Program Evaluation Forms and Approaches, Allen & Unwin, St. Leonards, NSW.

Parlett, M. & Hamilton, D. (1972) 'Evaluation as illumination: a new approach to the study of innovatory programs', University of Edinburgh, Centre for Research in the Educational Sciences, 1972. (Occasional Paper, 9).

Patton, M.Q. (1986) Utilization-Focused Evaluation 2nd Ed. Sage Publications, Newbury Park, California.

Patton, M.Q. (1987) When To Use Qualitative Methods In Evaluation, Sage Publications, Newbury Park, California.

Patton, M.Q. (1988) 'Integrating Evaluation into a Program for Increased Utility and Cost-Effectiveness', in Evaluation Utilization, (ed.) McLaughlin, J.A., Weber, L.J., Covert, R.W., & Ingle, R.G. New Directions for Program Evaluation, No.39 Jossey-Bass Inc. San Francisco, London.

Wadsworth, Y. (1991) Everyday Evaluation On The Run, Action Research Issues Association (Inc.), Melbourne.

Further Reading

Brown, L. (1988) 'Group Self-Evaluation - Learning for Improvement', School Improvement Plan Secretariat, Ministry of Education, Victoria.

Guba, E.G. (1978) 'Toward a Methodology of Naturalistic Inquiry in Educational Evaluation', Centre for the Study of Evaluation (CSE), University of California, Los Angeles.

Hamilton, D., MacDonald, B., King, C., Jenkins, D., & Parlett, M. (ed.) (1977) Beyond the Numbers Game, Macmillan Education Ltd. U.K.

Kemmis, S. & McTaggart, R (ed.) (1988) Action Research Planner, Deakin University

FIGURE 5.1: PROGRAM IMPROVEMENT - QUESTIONS TO ASK

	QUESTIONS TO ASK	FOCUS OF EVALUATION	KEY AUDIENCES FOR EVALUATION
NATIONAL	Are the national strategies designed to address literacy needs working? Are these areas in need of improvement? How effectively is the present infrastructure supporting the needs of the States? Are the personnel management policies appropriate or should there be changes? What are the strengths and weaknesses of the professional development delivery strategies? Are there further areas of professional development which are required to improve state, regional and provider operations? How are policies on resource allocation/distribution and curriculum affecting the states/territories, regions and providers?	Processes	Federal Ministers Portfolio/departments.
STATE/TERRITORY	How are current policies affecting the regions? Is the present system of allocating and distributing resources to individual regions appropriate, or should there be changes in priorities to improve overall service provision? How well is service provision co-ordinated across the State/Territory? Are there improvements needed in the types of professional development opportunities provided? Are the accreditation structures working? Are they sufficiently flexible and appropriate for diverse learner contexts and needs? How is the accreditation process implemented?	Processes	State/Territory planners Funding bodies
REGIONAL	How are present policies affecting providers? Are changes required in priorities for funding, professional development, or reallocation of resources? How are providers responding to the availability of curriculum development funds? What difference is the funding making to the quality of program delivery in the region? What are the communication linkages between providers within and between regions? Could these be improved?	Processes	State and Regional Authorities Funding bodies
PROVIDER	What happens in the program? What are practitioners doing that is working well? What is not working so well? How are students affected by the program in action? Is the curriculum delivery tailored to meet individual needs and goals of students? Are professional development strategies effective as identified by staff? How has professional development affected curriculum provision? Are structures for program evaluation in place? Are they used?	Processes	Program managers Program staff and clients

FIGURE 5.2: PROGRAM COMPONENTS: REVIEWING SERVICE PROVISION

- Philosophy
- Planning
- Community Linkages
- Awareness Activities - Recruitment, Public Relations, Advocacy
- Accessibility of Provision to Potential Clients
- Facilities and Equipment
- Administration
- Participation
- Staff Development and Training
- Tutor Training
- Volunteer Tutor Support
- Learner Assessment
- Learner Support Services
- Instructional Strategies
- Materials
- Evaluation Methods
- Funding

Source: Components drawn from 'Good Practice Statement', Adult Literacy Volunteer Tutor Program Evaluation Kit, Ministry of Advanced Education and Job Training and Ministry Responsible for Science and Technology, Provincial Curriculum Publications, British Columbia, 1989.

316

317

Chapter 6 EVALUATION FOR PROGRAM JUSTIFICATION/ ACCOUNTABILITY

Broadly, a major purpose of undertaking evaluation is to account for the resources expended on the program under review. Generally, accountability evaluation fulfills the need for someone directly associated with the program to report to someone else, for example, for a program manager in government to report to Treasury. In this scenario, the manager justifies the funds and other resources used to provide the program.

Within the ALBE context, the National Framework of Adult English Language, Literacy and Numeracy Competence could be used as the basis for setting up a monitoring or justificatory evaluation.

For example, the effectiveness of a given provider could be monitored using simple measures of effective completion of their adult literacy programs across all aspects of competency.

Alternatively, the worth of a given program could be judged according to how effective it was in moving students towards independence on one or more of the key aspects of competence.

A distinction is made here between two approaches to evaluation for program justification and accountability:

- the *monitoring* of both the delivery and outcomes of programs that are well-established. The purpose of evaluation is to improve the internal management of programs and to provide the basis for enhanced external accountability. Evaluation is one element of a cyclical process which includes program development, resource allocation and program delivery.

- *impact evaluation* which also focuses on the delivery and outcomes of well-established programs, but where the emphasis is on the assessment of outcomes. The primary purpose of evaluation is to justify the worth of the program and to account for program expenditure.

Monitoring Evaluation

Monitoring evaluation is an integral part of the management of programs. Managers need to have information that will help them plan and control service delivery as well as to account for their program and organisational responsibilities.

Project management consists of:

- a strategic plan for the implementation of policy
- a program structure which reflects the strategic plan and which has goals and objectives stated in terms of desired outcomes
- management arrangements to ensure implementation of the program structure (and, in some instances, a data-base on facets of program provision i.e. a management information system or systems (MIS))
- the use of the MIS for making decisions about program policy, monitoring program delivery and assessing outcomes (Owen, 1993:110).

Managers might use monitoring evaluation to answer such questions as:

- Is the target population being reached?
- Is the program being implemented in the ways specified?
- Is the program effective?
- How much does it cost?
- What is the cost relative to its effectiveness?

Figure 6.1 lists questions that might be asked within monitoring evaluation at different levels within ALBE operations.

The emphasis is to use monitoring evaluation to aid **performance oriented program management.** There has been a shift in focus from monitoring performance in terms of dollars spent or number of clients assisted, to value for money and outcomes achieved through public spending (Wholey & Hatry, 1992:604). Performance monitoring is thus a critical component of accountability as well as a tool for internal performance management.

Monitoring to provide performance information is simply an 'organisational method of learning from experience' and involves the use of timely feedback for continuous adjustment of policies and practices. This 'evolutionary' approach to development is presented in the program management cycle given in Figure 6.2 (Stuart & Powrie, 1993).

In Chapter 4, we described in detail a way of structuring performance information for evaluation purposes. Figure 4.6 shows how program logic statements, and in particular, an expanded objectives hierarchy model can be designed to provide the basis for both monitoring and impact evaluation. The case study example demonstrating the objectives hierarchy model for a workplace basic education project is repeated here in Figures 6.3 and 6.4. Figure 6.3 shows the simple hierarchy of objectives and their underlying assumptions, while Figure 6.4 is a plan which connects objectives, implementation strategies and performance information. Fundamental to the hierarchy of objectives is the assumption that the attainment of the first objective is a prerequisite for the achievement of the next objective in the chain. Planning a program using this kind of model has considerable merit. It requires program managers (in consultation with stakeholders) to think about the results they want to achieve in a cause-effect chain and to consider what they should be measuring as the program develops over time. Identifying the performance information required for each objective provides the basis for considering which objectives (and which program components) are more amenable to monitoring than others. Once decisions about performance information are made, reporting requirements can be structured accordingly.

Monitoring Performance

Performance monitoring differs from collecting program statistics in that it implies some 'standard of comparison' and some sense of desired direction (Wholey, 1993). The following standards of comparisons are those typically used in the measurement of performance of public programs:

- comparisons with prior performance
- comparisons with the performance achieved in similar programs
- comparisons with target levels of expected performance competency.

Performance measures include:

- "Raw" measures of program outcome (e.g. numbers of program graduates in re-training programs)
- Outcome measures that are adjusted to reflect "degree of difficulty" related to differences in client characteristics or other local conditions that influence program outcomes
- Gain scores (or changes in outcomes e.g. increases in numbers of program graduates in re-training programs)
- Measures of program impact (the difference between program outcomes and the outcomes that would have occurred in the absence of the program (Wholey, 1993:7).

Performance monitoring should draw upon the full range of information collection and analysis techniques and include both qualitative and quantitative data. However, with public agencies coming under increasing pressure to demonstrate efficiency and effectiveness in all their operations, there has been a strong emphasis on the use of performance indicators to monitor programs in order to make decisions about how well they are working.

Performance Indicators and their Uses in Monitoring Programs

A performance indicator is a simple statistic designed to give summary information which allows managers and others to make judgements about the effectiveness of programs. Information is usually expressed as an index, rate or other numerical figure. Performance indicators are, in effect, substitutes for things that cannot be assessed directly. For example, a medical symptom such as 'a raised temperature' is an indicator of illness (NSW Public Service Board, 1986).

'In themselves performance indicators do not show whether a program's performance is satisfactory. They usually have to be compared with some 'base figure' - for example, with a target, with a national or international standard, with performance of other components of the same program or with the performance of the program at a different point in time.' (Social Justice Strategy Unit, 1988:5)

Within ALBE operations, performance indicators might be used at all levels in the decision-making process. For example, they might be used by:

- **program staff,** in focussing their efforts on program impact, especially when they have been actively involved in setting the goals and the indicators. 'The availability of reliable and timely performance information opens the way for extended delegation of authority, allowing people to take greater control of their jobs and moving decision-making closer to the workface. This can lead to improved productivity as staff are stimulated to think creatively about the way performance objectives can be achieved and to become more concerned with the quality of the services they provide.' (Program Strategic Review Branch, 1990:13). (It should be noted that the performance indicators have the potential to influence programs in unintended and undesirable ways - see the discussion below on validity and ethical issues in the use of performance indicators.)

- **individual program managers and service providers,** in setting standards of performance and targets, identifying program impact and improving effectiveness. Performance information enables managers to critically examine their goals, strategies and processes and provides the basis for discussion with staff and clients about program improvement.

  Performance indicators are also becoming increasingly important for the credibility of bids for resources as part of the annual budget process. Given that performance information can potentially be used by funding bodies to scale down or terminate program activities, it is not surprising that the use of performance indicators is met with some anxiety and antagonism (refer below to 'Encouraging Monitoring In Program Management).

- **regional, state and national authorities,** in reviewing performance against broad goals, assessing the continued appropriateness of those goals and revising existing strategies and priorities accordingly. Performance indicators provide input to the strategic planning process, decisions on program structures, organisational design and resource allocation.

- **ministers and department heads,** for accountability purposes and allocation of resources within the budget. Performance information will also be useful to them in their dealings with client groups and their advocates and in presenting their policies to the public.

- **program clients,** for developing informed views about programs that are of concern to them, enabling them to contribute more effectively to debate about program policies and practices.

Monitoring In Practice: Examples

(a) Case Study A: Post-Program Monitoring (PPM): Labour Market Programs

Figure 6.5 is a description of post-program monitoring (PPM) which has been administered since 1988 by the Commonwealth Department of Employment, Education and Training (DEET). Its purpose was to contribute to efficient spending on labour market programs which are characterised by a single enduring objective of assisting 'clients to obtain jobs or to move into further education or training following their program assistance' (Stuart & Powrie, 1993: Supplementary Papers). It is not only used for program budgeting, but also for operational planning in that Area managers agree to targets for program outcomes as measured by the PPM system. PPM data is conveyed to Areas on a monthly basis through an on-line Management Reports System.

A collection of data dating back to late 1988 now exists. Based on this data, the effectiveness of a range of programs can be compared for given client types, and in varying labour market conditions. Evidence indicates that the system has effectively guided cross-program allocation decisions at the national level.

(b) Case Study B: National Management Information and Statistics Systems (NATMISS)

NATMISS is a project which aims at providing a national reporting system as a basis for the measurement of performance outcomes in the vocational education and training sector. The project will be operational in mid 1994 and is designed to achieve a range of objectives including:

- determining training demands and resource needs
- providing the basis for consistent reporting by all providers
- assisting with the planning, monitoring and provision of programs at all levels, and in particular, alternative learning and delivery strategies
- accurately identifying gaps in equity and access to programs
- informing the public and government of the achievements
- justifying additional funding.

What is NATMISS?

'The NATMISS Project can be viewed at two levels: As a database and as an executive information system.

The database will comprise data collected from student enrolment forms, contracts of training and other institutional records such as financial management systems and human resource management systems. This is quantitative data collected on a regular basis. It will give a picture of the preceding year's activities. Importantly it will provide an accurate assessment of outcomes, including module and course completions.

The NATMISS design will provide consistency of reporting by all states, and permit accurate international comparisons, through the extensive use of Australian and international standard classifications.

The NATMISS executive information system (EIS), when designed in the implementation phase of the project, will include the NATMISS data base and information from other sources including the Australian Bureau of Statistics, satisfaction surveys and authority reports. All of this information will be drawn together into a single integrated system through the use of the EIS operating in an open systems environment.

> ... By using the National Goals and Objectives to provide broad policy framework, agreed to by all States and Territories, all data elements identified are required to support identifiable needs. Inspection of the data model ... (See Figure 6.6)... shows that a great deal more information can also be produced which is relevant to vocational education and training but not restricted to that required by specific performance indicators. However, the performance indicators which can be derived from the NATMISS database are essential to the measurement of achievement of the participation, equity and reform objectives which have been nationally agreed to.' (S.T.A.T.S.,1993)

Types of Performance Monitoring

Performance indicators can be used as tools for four major types of program monitoring relevant to ALBE operations. These are:

- **effectiveness** i.e. the extent to which a program is satisfying the purpose for which it was established.

  For example:
  National Level Effectiveness Indicator:
  Population making literacy gains against national competence standards: total population.

- **social justice** i.e. the social impact of a program, in terms of equity (a fair distribution of resources among client groups), equality (equal civil, legal and industrial rights for all), access (equitable access to services) and participation (the opportunity for individuals to participate in the decisions which affect their lives). Social justice might also be viewed as a dimension of program effectiveness.

  For example:
  National Level Equity Indicator:
  Number of priority group enrolments (e.g. unemployed): total enrolments.

- **efficiency** i.e. the relationship between inputs and outputs. Efficiency indicators are concerned with productivity (e.g. changes in unit costs), and aspects of service delivery (e.g. the ratio of successful to total applications for a service).

  For example:
  National Level Efficiency Indicator:
  Cost of curriculum delivery: successful outcomes.

- **Cost-effectiveness** i.e. the unit cost of outcome. Effectiveness measures alone take no account of the resources consumed to achieve objectives. When considered together to measure overall performance, efficiency and effectiveness measures are referred to as cost-effectiveness indicators (Social Justice Strategy Unit, 1988: 7-8).

  For example:
  Provider Level Cost-Effectiveness Indicator:
  Program costs related to the number of employees who had completed the workplace basic education course: the number of employees who had gone on to participate in industry training programs.

Principles for Developing Performance Indicators

Performance indicators should meet certain standards of technical quality and be useful.

Technical quality will depend on the following criteria being met:

- **Validity** i.e. does the indicator measure a program outcome which can be attributed solely to that program?

- **Accuracy** (reliability) i.e. is each indicator a true measure of the index intended?
- **Timeliness** i.e. is the information likely to be available when needed?
- **Cost-efficiency** i.e. can the information required be collected at a reasonable cost?

To be useful, performance indicators need to be:

- linked to program objectives and management processes
- relevant to decisions about program improvement
- available in a user-friendly format to local decision-makers as well as to the central program managers to whom they are accountable (Hatry, 1990).

Each indicator and the indicators as a set should be checked against the following criteria:

- **Significance** - does each indicator measure some aspect of the program which is of significance? Are the data worth collecting?
- **Uniqueness** - does each indicator provide information not provided by any other in the set, or in any other way?
- **Comprehensiveness** - do the indicators, taken as a set, cover the objectives as a whole?
- **Norms** - are the norms on which judgements are made about the program clear ? Will the indicators be used to compare measures (i) from the same program at other times, (ii) from other programs at the same time, or (iii) with an absolute standard ?
- **Interpretability** - will decision makers be able to make sense of the information? (Owen, 1993:126-127)

A sample set of indicators which ALBE staff might use to monitor their programs is provided later in this chapter (see p.52). However, there may be a need for providers and others to develop indicators for a unique situation. Figure 6.7 is a worksheet which provides a structured approach to the development and implementation of performance indicators. This approach explicitly addresses the non-rational or political issues which surround the use of performance indicators.

Encouraging Monitoring in Program Management

When used for accountability purposes, monitoring can be threatening to managers and staff whose programs are being evaluated. Wholey makes a range of suggestions for encouraging monitoring evaluation in program management. Among these are:

- creating incentives for the constructive use of evaluation.

  Firstly, by encouraging evaluation effort through the requirement to monitor on a regular basis and report on program outcomes.

  Secondly, by 'rewarding' managers and staff for identifying opportunities to improve efficiency. For example, adopting a policy which provides for the sharing of 'savings' achieved from improved service quality that results in lower costs, rather than removing all the savings.

- involving managers and staff in program evaluation efforts i.e. using program personnel in helping to identify program objectives and evaluation criteria (Wholey, 1993:3-222).

Validity and Ethical Issues in the Use of Performance Indicators

Strong warning needs to made against relying exclusively on performance indicators in monitoring programs. Indicators may reflect only a part of the program reality and those aspects which are:

- easily measured
- of particular concern to funding bodies, for example, outcomes
- common to a number of programs and therefore useful for comparisons (Batterham, 1991: 5-6).

This has been described as the 'reductionist threat' to the validity of findings based on information generated on indicators alone. Total reliance on performance indicators inevitably leads to bias, a failure to capture the richness of the program, and may provide a disincentive for development and innovation. Monitoring based exclusively on indicators may ignore:

- aspects of the program which present measurement difficulties or are best captured by description as opposed to measurement
- concerns of program providers and their clients
- goals of program providers and their clients
- elements of the program which are unique or innovative
- program processes (Batterham, 1991:6).

Extreme care must be taken to ensure that the interpretation of indicators is carried out with an awareness of their limitations in representing program reality. Added to this, the '**validity of the interpretation will depend directly on the quality of qualitative data available to inform the interpretation.**' (Batterham, 1991:7)

Those who plan to use performance indicators in program monitoring should be aware that they may have unintended consequences:

- Where attempts are made to establish centrally determined indicators for the purpose of monitoring and comparison, indicators tend to promote conformity by inhibiting innovation and goal expansion (Singh, 1988: 3-5).

- Indicator data can give misleading information about what is significant and its relative importance. 'For instance, the appearance of poor performance may conceal the fact that social justice values are given preference over instrumental values of "efficiency" and "effectiveness".'(Singh, 1989:3)

- There is the danger that the indicators can become 'ends in themselves, driving instructional practices in unintended directions, irrespective of whether or not they are used by policy makers (David, 1988, in Singh, 1988). Batterham (1991:16) suggests that 'indicators, rather than being a path to good practise, end up defining good practise albeit in a very reductionist and minimalistic way. This leads to practises like teaching to the test in education...' This is referred to as goal or objective displacement.

- The tendency to use indicators related to outcomes that can be measured in the short-term may well lead service providers to accept that these are the outcomes that policymakers value and that are important while significant longer-term outcomes may be ignored.

- Indicators may not provide evidence about outcomes which can only be described, nor about those outcomes which are unintended but nevertheless important.

- The evidence suggests that centrally generated, uniformly applied indicator sets can be inaccurately interpreted. This might be explained by a lack of understanding of the context in which the performance measurement is made, or be related to the political purposes being served by the information. Skilbeck, when considering the question of accountability in schools, has suggested that:

'Well constructed individual school monitoring procedures, accompanied by regional moderation and an agreed upon reporting procedure acceptable to the community at large, may

> be an acceptable alternative to state or nation-wide assessment through the use of standardized tests (Skilbeck, 1980 in Singh, 1988).

Those at all levels within ALBE operations who are concerned with monitoring should be aware of the potential effects of relying too heavily on indicators alone as a basis for program justification and accountability.

D'Cruz (1989/1990:11-12) argues strongly for policymakers to think carefully about the meaning of accountability in the context of adult education given its complex objectives and activities. He suggests that accountability should not only include those outcomes of provision that are readily quantified but should be associated with additional notions such as 'explaining', 'describing', 'making clear', and justifying one's intended value criteria and programs of action.

Sample Performance Indicators

The previous sections have outlined key attributes of performance indicators and some concerns about the exclusive use of indicators within evaluation. It should be emphasised that indicators are indicators, that is they need to be used to assist managers and others make sense of a given situation. Valid conclusions about the effects of a program can only be drawn where there is an understanding of the program context in which the performance information was collected.

The purpose of this section is twofold. First, it is to show how indicators are linked to specific program objectives and second, it is to provide some samples of indicators which might be employed in a program evaluation situation. The sample indicators are presented in Figures 6.8 to 6.27.

The indicators are grouped according to the level of operation at which an evaluation might be planned. The levels of operation are:

- national Figures 6.8 to 6.12
- State/Territory Figures 6.13 to 6.17
- regional Figures 6.18 to 6.22
- provider level Figures 6.23 to 6.27

which is consistent with the classification of questions to ask in a monitoring evaluation, which is outlined in Figure 6.1.

At the *national* level, indicators are presented as follows:

Figure	Operational Function
6.8	Policy and Planning
6.9	Program Development and Implementation
6.10	Personnel Management and Professional Development
6.11	Curriculum
6.12	Information Services and Financial/Resource Management

A similar suite of indicators are presented for the State/Territory, regional and provider levels using the same operational functions.

All of these tables have a similar format. Objectives relating to the operational function have been prepared and for each objective, a set of indicators are listed. It is important to note that each indicator involves a statement of the information which needs to be collected. In some cases, it involves more than one statistic and the manipulation of these statistics by some simple formula. It should be stressed that the use of these indicators implies that a sound data base of basic statistical information is available. In some instances the derivation of an indicator relies on the manipulation of data collected in a 'raw' form, for example on a questionnnaire return. Collection of raw data is discussed in the following two chapters of these guidelines.

Impact Evaluation

Impact evaluations are concerned with collecting information about program outcomes. In contrast to monitoring evaluation which is a continuous process, impact evaluations are undertaken as an 'end of program' activity.

Key questions in impact evaluation are:

- What are the 'real goals' of the program and to what extent have they been attained?
- What are the intended and unintended outcomes of the program?
- How do differences in implementing the program affect the outcomes?
- Does the program meet identifiable needs?

Figure 6.28 provides examples of impact evaluation questions which might relate to each of the levels of operation within ALBE.

Morris et.al. (1987: 21-31) provide some guidelines for developing performance measures. It is important to:

- determine the use that will be made of the findings by different audiences

- develop and clarify program objectives. Objectives should:
    - be clearly stated
    - be actively pursued by program activities (i.e. they are not just 'official' objectives stated in program documents)
    - reflect the level of skill/performance which the program as implemented is designed to develop
    - be organised to reflect their relative importance so that those of highest priority can be identified.

- establish whether program objectives are likely to be accomplished within the time-frame of the evaluation so that decisions can be made about whether it is appropriate to focus on outcomes or processes or both.

- decide which objectives can be measured by existing measurement instruments and which will require specially tailored instruments.

Three Approaches to Outcomes Evaluation

Morell (1979:2-3) argues that there is value taking a liberal interpretation of 'outcome' evaluation in the social service context. Based on a survey of the literature, he identifies three 'impact' evaluation types:

- **client comparison** studies, which involve studying the different effects of a program on different client groups
- **follow-up studies**, which require following up those clients who have left the program
- **modality tests** which attempt to judge the extent to which a program has brought about desired change.

Strictly speaking, it could be legitimately claimed that 'modality' tests are the only pure outcome evaluations in that they are the most direct tests of program influence. Follow-up evaluation deals with program effects, but in a diffuse way in that it is concerned with the complex interaction between program/intervention and the life situation of the individual. Client comparisons focus on the individuals receiving the intervention as opposed to the intervention itself. However, both client comparison and follow-up studies are included here because, as highlighted below, both can contribute useful information related to program impact.

(i) Client Comparison Evaluation

Client comparison evaluation provides information needed for the rational planning of services. It is used to:

- identify those individuals or groups who are most likely to be helped by a program. This information can become the basis for improving the quality of service targeted at these groups.

- determine which groups stand to benefit most from particular programs (i.e. for screening purposes). It may well be that screening is not used for a variety of ethical, philosophical or political reasons for as Etzioni (1969) has illustrated, organisations are not singularly devoted to maximum efficiency in the use of their resources. Regardless, appropriate allocation of clients to particular services is a critical part of program planning. Client comparison evaluation has an important contribution to make in weighing up legitimate arguments against excluding people from services and arguments for planning based on a realistic assessment of whether people can be helped.

- understand a program's pattern of successes and failures i.e. a knowledge of the characteristics of specific client groups may well reveal why some groups benefit more than others (Morell, 1979:10-11).

(ii) Follow-Up Studies

A judgement about program influence will require follow-up information concerning:

- the extent of the beneficial effects and how long they last (attenuation effects)

- any unintended consequences, beneficial or otherwise. Human service programs impact their clients' lives in numerous ways, many of which cannot be predicted prior to program implementation

- the dynamics by which program effects change over time (Morell, 1979:14-15).

In many cases, follow-up evaluation is the only form of evaluation which can be carried out. This may occur when the program has relatively little control over the people it must accept, if it does not have the potential to do sophisticated screening, and if the existence of the program is a foregone conclusion due to political or other circumstances (Morell, 1979:16).

(iii) Modality Tests

Modality tests are concerned with how people are being helped and are 'based on the assumption that a particular treatment (or some part of it) will result in well-defined, detectable changes in clients.' (Morell, 1979:19) Modality test evaluation can be used:

- to indicate the extent of changes resulting from an intervention
- in combination with process evaluation information to produce an in-depth understanding of the program or intervention process
- to determine which aspects of the program led to identified outcomes.

Impact evaluations using modality tests require that program staff define their goals and objectives. Many reasons have been given why this task is often challenging, for example:

- difficulties associated with articulating motivations
- general feelings of apprehension about the evaluation and its consequences

- the fact that social service organisations have diffuse, obscure, and continually shifting goals (Patton, 1978).

Nevertheless, it is critical that the 'real goals' of a program are identified, namely those goals which are reflected in program practice.

Sometimes the translation of goals into measurable outcomes and the analysis of findings can involve complicated statistical procedures. It is therefore, in this area of outcomes evaluation, that the assistance of external evaluation expertise would appear most critical. Difficult data-management decisions often relate to limitations on data sources, funding and time. In many instances, compromise is necessary and surrogate measures which replace ideal measures, must be used.

Justification and Accountability Decision Making

Ultimately, evaluations described in this chapter lead to the following kinds of decisions:

- Is the funding allocated to the program justified?
- Can the public have confidence in the services offered through the program?
- Can praise or blame for the success or failure of the program be allocated to program personnel?
- Should the program be discontinued because of its failure to provide adequate levels of outcomes? (Owen, 1993: 115)

References

Batterham, R. (1991) 'Are clinical indicators bullies?' Paper available from the Centre for Program Evaluation, The University of Melbourne (mimeo). June, 1991.

D'Cruz, J.V. (1990) The D'Cruz Report, Adult Basic Education and the Public Culture of Victoria, Division of Further Education, Ministry of Education, Melbourne, Australia.

Etzioni, A. (1969) Two approaches to organizational analysis: A critique and a suggestion. In Program Evaluation in the Health Fields, ed. H.C.Schulberg, A. Sheldon and F. Baker, New York: Behavioral Publications.

Hatry, H.P.(1990) 'Performance Measurement: A Major Management Tool', Conference Proceedings, of the National Evaluation Conference 1990, Vol.1. ISBN 0 7305 7812 7, 0 7305 7813 5.

Morris, L.L., Fitzgibbon, C.T. & Lindheim, E. (1987) How to Measure Performance and Use Tests, Sage Publications, Newbury Park, California.

Morell, J.A. (1979) Program Evaluation In Social Research, Pergamon Press Inc. New York.

NATMISS. An Information Blueprint for the 21st Century, (1993) VEETAC Committee on TAFE and Training Statistics.

Owen, J.M. (1993) Program Evaluation, Forms and Approaches, Allen Unwin Pty. Ltd., St. Leonards, NSW.

Program and Strategic Review Branch (1990) 'Guidelines for the Development of Performance Indicators', Queensland Treasury, September, 1990.

Singh, M.G. (1988) Some Issues in the Use of Performance Indicators in Education. A Working Paper presented for discussion to the Annual Conference of the Australian Association for Research in Education, University of New England, Armidale, NSW, 30 Nov.- 4 Dec., 1988.

Singh, M.G. (1989) 'Measurement of Progress: "Kill Ratio", "Body Counts" and Other Performance Indicators, The Australian Administrator, School of Education, Deakin University, October 1989.

Social Justice Strategy Unit (1988) 'Performance Indicators and the Social Justice Strategy', Discussion Paper, Department of Premier and Cabinet, Victoria, August.

S.T.A.T.S. (1993) 'Vocational Education - The Force of the Nineties', The VEETAC Committee on TAFE and Training Statistics Newsletter, May, Issue 4.

Stuart, A. & Powrie, S. (1993) 'Performance Monitoring in DEET - A Model for Performance Feedback', Supplementary Papers, The Australasian Evaluation Society International Conference, 1993.

Wholey, J.S. & Hatry, H.P. (1992) 'The Case for Performance Monitoring', in Public Administration Review, November/December, Vol.52, No. 6.

Further Reading

Community Services of Victoria. (1990) 'Guidelines for Developing Performance Indicators', Draft, March.

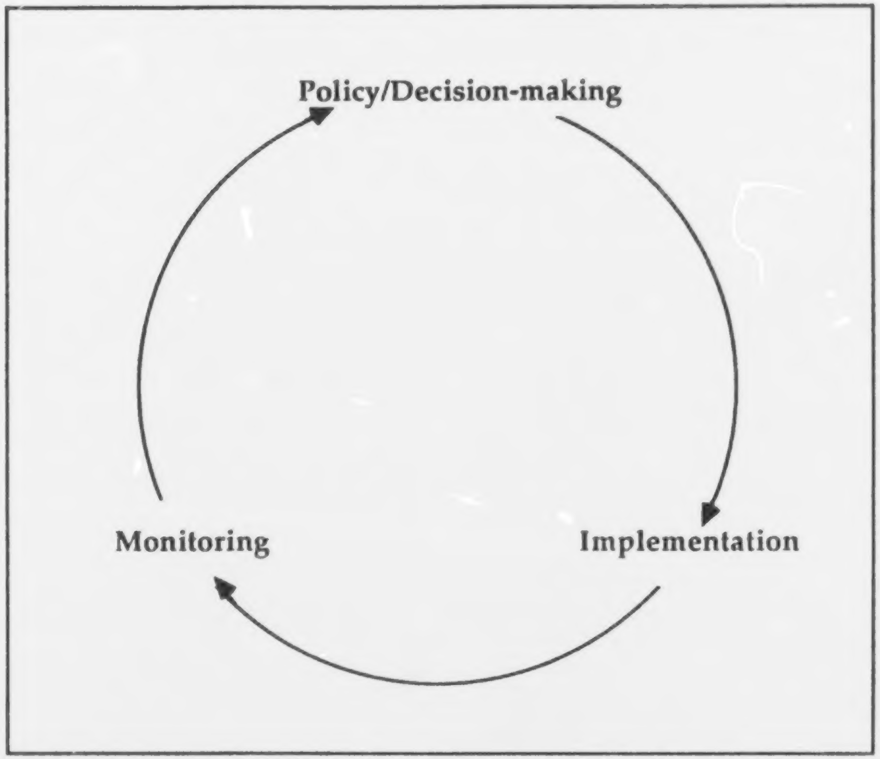
Guthrie, H. (1988) 'Performance Indicators in TAFE, TAFE National Centre for Research and Development, Adelaide.

Wyatt, T., Ruby, S., Norton, B. Davies & Shrubb, S. (1989) 'Reporting on Educational Progress: performance indicators in education', A Report to the Conference of Directors-General of Education, Sydney.

FIGURE 6.1: MONITORING EVALUATION - QUESTIONS TO ASK

	QUESTIONS TO ASK	FOCUS OF EVALUATION	KEY AUDIENCES FOR EVALUATION
NATIONAL	How many people have participated in the programs? Have literacy skills increased? What has been the effect of increased Commonwealth expenditure? Has the money allocated to strategic areas been expended in these areas? What was the cost of identified outcomes? Could the same outcomes be achieved with less inputs? What impact has curriculum development funding had on outcomes? What difference has increased funding made to the type of student outcomes? How much do we spend on each student? What impact has expenditure had on curriculum and professional development? What impact is professional development having on the teaching service and program?	Existing Programs	Federal and State Ministers Finance and Treasury, Portfolio/departments.
STATE/TERRITORY	Are state policies being implemented? Are priority areas being provided for? What impact has curriculum development funding had on outcomes? Have skill levels improved? What difference has increased funding made to the type of student outcomes? What impact has expenditure had on curriculum and professional development?	Existing Programs	State/Territory planners Funding bodies
REGIONAL	What proportion of teaching staff are funded according to award and benchmark conditions? How has the regional and professional development policy been implemented? What effect is it having on the quality of programs offered in the region? How are providers responding to the availability of curriculum development funds? What difference is the funding making to the quality of program in the region? Has there been an impact on achievement of student learning outcomes? How effectively have regional funds been used?	Existing Programs	State and Regional Authorities Funding Agencies
PROVIDER	How many students are achieving success in our programs? What is the attrition rate from our programs? Are our clients getting jobs or improving job options? What proportion of our students are going on to further studies? Do our programs exemplify national/state good practice? What impact is professional development work having on programs? How much does it cost to run the programs? How efficient are programs in producing desired outcomes?	Existing Programs	Policymakers Funding agencies Program co-ordinators and staff.

FIGURE 6.2: PROJECT MANAGEMENT CYCLE


Policy/Decision-making
Implementation
Monitoring

Source: Stuart, A & Powrie, S. 'Performance Monitoring in DEET - A Model for Performance Feedback', Supplementary Papers, AES Conference, June 1993.

FIGURE 6.3: AN OBJECTIVES HIERARCHY FOR A PROPOSED WORKPLACE BASIC EDUCATION PROGRAM

Hierarchy of Objectives	Underlying Assumptions
III. Ultimate Objectives	
9. To meet the literacy needs of adults - needs which relate both to their work and other life experiences.	Adults whose literacy needs are met will lead fuller, more satisfying lives.
8. To maintain and improve national productivity by raising literacy skills in the workforce.	The modern industrial setting with its rapidly changing technology demands a flexible and educated workforce. The acquisition of literacy skills is essential for workplace efficiency and the re-training of employees.
7. To facilitate and legitimise the establishment and maintenance of workplace literacy programs.	Planning bodies and employees will support workplace literacy programs once their benefits have been demonstrated.
II. Intermediate Objectives	
6. To enhance employee self-confidence and interpersonal skills.	Self-confidence and interpersonal skills are an integral part of the effective use of literacy skills in the workplace.
5. To improve employee morale.	Employees will respond positively to organisations that provide on-site literacy programs which meet their needs.
4. To provide programs that meet the literacy needs of employees in their specific workplace contexts.	Employee literacy skills are most effectively addressed via on-site programs tailored to their needs in the specific workplace context.
I. Immediate Objectives	
3. To train practitioners to undertake literacy programs designed specifically for workplace contexts.	Workplace literacy programs are most effective when they: • involve management, unions and literacy provider in the setting up process • are custom built following a careful literacy audit and training needs analysis • are presented as part of an overall workplace training strategy rather than a remedial service • balance vocational skills development with the skills necessary for broader social interaction • become embedded as an ongoing feature of the workplace training strategy. (Ref: Workbase NZ outline for Literacy Audit)
2. To delineate organisational/individual employee learning needs.	
1. To create an awareness of and support for the creation of on-site literacy acquisition programs amongst employers, unions and potential clients (employees).	

FIGURE 6.4: AN EXPANDED OBJECTIVES HIERARCHY FOR A PROPOSED WORKPLACE BASIC EDUCATION PROGRAM

Hierarchy of Objectives	Strategies/Activities	Performance Information
II. Intermediate Objectives		
6. To enhance employee self-confidence and interpersonal skills.	Conduct individual and group learning activities designed to foster confidence and interpersonal skills.	Observations of changes in employee behaviour made by managers/supervisors
5. To improve employee morale.	Provide positive feedback.	Reduction in staff turnover.
4. To provide programs that meet the literacy needs of employees in their specific workplace contexts.	Tailor program content and activities to address the immediate literacy needs of employees.	Assessment of student reading/writing/mathematical skills on work related tasks. Course completion rate. Attendance rate. Student satisfaction rating. No. of student hours required for course completion.
I. Immediate Objectives		
3. To train practitioners to undertake literacy programs designed specifically for workplace contexts.	Conduct workshop sessions to train tutors to conduct literacy programs designed to meet needs identified in specific workplace contexts.	Level of staff satisfaction Course completion rate Attendance rate
2. To delineate organisational/individual employee learning needs.	Examine company job task breakdown sheets, observe workers at work, record reading/writing tasks undertaken on the job, collect reading/writing material which pertains to the job, interview potential participants to assess personal literacy learning needs, etc.	Student rating of program Employer satisfaction rating Course completion rate Attendance rate
1. To create an awareness of and support for the creation of on-site literacy acquisition programs amongst employers, unions and potential clients (employees).	Discussions with management/unions/employees regarding benefits of on-site workplace literacy programs.	Number of organisations willing to participate Employee enrolment: places offered Enrolments: applicants Gender ratio Penetration rates (e.g. of designated priority groups)

FIGURE 6.5: POST-PROGRAM MONITORING SYSTEM (PPM) LABOUR MARKET PROGRAMS (DEET)

How PPM Works

The PPM survey collects information on the job, training and education outcomes of clients who have received labour market program assistance.

The data are collected over a period of 6 weeks by an initial mail survey and, if necessary, a reminder mail survey and telephone survey. Each fortnight, the PPM system selected those clients whose program assistance ended in a two week period (the survey fortnight) about 9 weeks ago. A check of Departmental program databases is then made to see whether the clients selected for PPM follow-up have subsequently been given further program assistance, and a PPM outcome may be computer-generated to indicate the type of further program assistance undertaken. Clients who have not been given further assistance are sent a PPM questionnaire. Clients return their completed survey form in a reply paid envelope enclosed with the questionnaire.

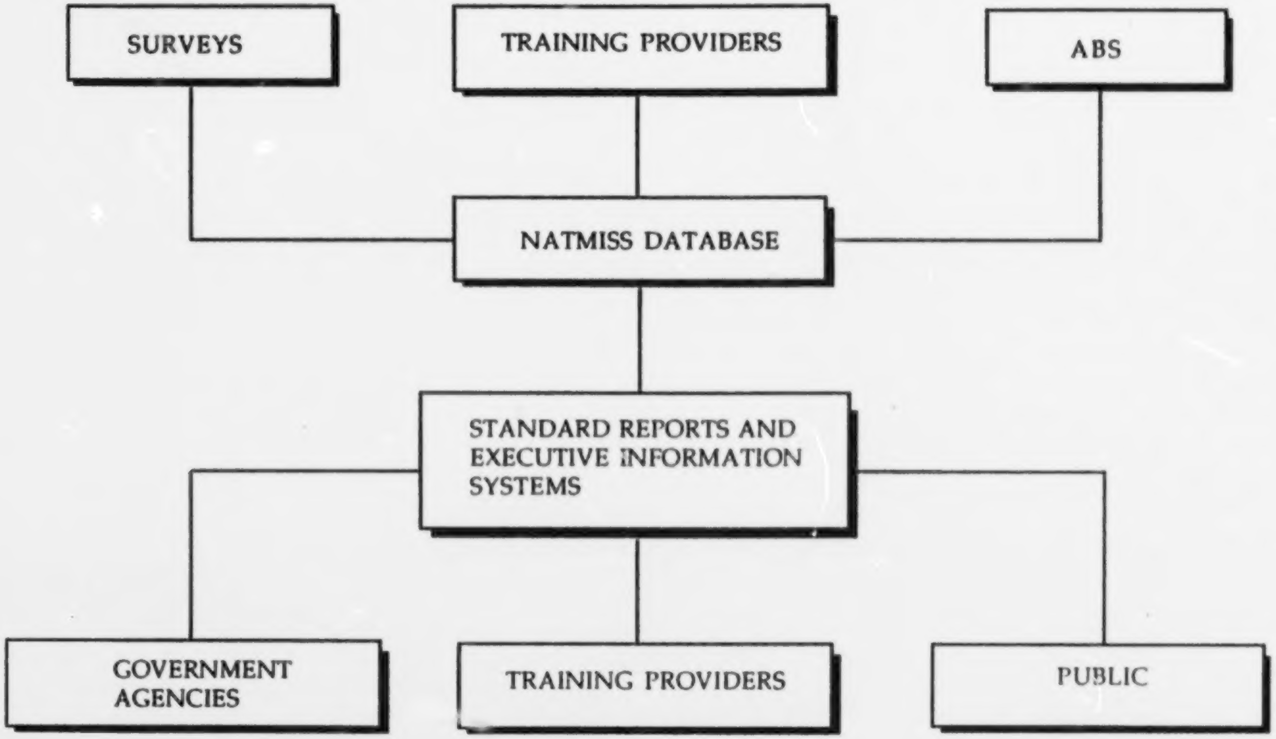
Three weeks after the first mail survey is despatched, clients who have not returned the first PPM questionnaire and who are not in further program assistance are sent a reminder questionnaire. Approximately 2 weeks after the reminder mail survey is despatched a telephone survey is begun of those clients who have not returned their PPM questionnaire.

Some 16 weeks after the clients ceased assistance, the PPM follow-up process is complete and data are available for reporting.

Source: Stuart, A. & Powrie, S. 'Performance Monitoring in DEET - A Model for Performance Feedback', Supplementary Papers, AES International Conference, July, 1993.

FIGURE 6.6: NATMISS INFORMATION FLOW

NATMISS mersures inputs to, outputs from and outcomes of Vocational Educational and Training. NATMISS relies on the information collected from training providers, through special surveys and from the Australian Bureau of Statistics. This information is processed by NATMISS and reported in a concise manner back to training providers.


SURVEYS
TRAINING PROVIDERS
ABS
NATMISS DATABASE
STANDARD REPORTS AND EXECUTIVE INFORMATION SYSTEMS
GOVERNMENT AGENCIES
TRAINING PROVIDERS
PUBLIC

Source: S.T.A.T.S. (1993) 'Vocational Education - The Force of the Nineties', The VEETAC Committee on TAFE and Training Statistics Newsletter, May, Issue 4.

FIGURE 6.7: A STRUCTURED APPROACH TO THE DEVELOPMENT AND IMPLEMENTATION OF PERFORMANCE INDICATORS

The aim of this worksheet is to present, in a reasonably structured way, the main issues for consideration in the development and implementation of performance indicators. It is by no means the only way and is certainly not intended as a constraint on the desire of interested groups or individuals to pursue different approaches.

1. What are your unit's goals, objectives, expected outputs, customer needs and critical success factors ?

2. What information would enable you and your staff to judge the calibre of your unit's performance on a regular basis ?

3. Sketch your understanding of the information your co-ordinator/team leader and then his/her manager would find valuable in assessing your work unit's performance.
    - Your co-ordinator/team leader
    - His/her manager

4. To what extent is there an alignment between your performance information (see Question 2) and that expected by/valuable to the two managers documented above ?

5. Now undertake a similar process in regard to client needs for information.
    - How would you know if student/client needs were, in fact, being met by the services you deliver ?
    - What information would you publish in your Annual Report or Budget documents to show how your program was achieving Government policy objectives and satisfying student/client needs?

6. Return to your own performance information in Question 2 and revise it in the light of the additional judgements you have made in Questions 3, 4 & 5.

7. Divide your performance information now into the following groups:
    - Those indicators easy to collect information about with little extra effort.
    - Those indicators most valuable in accurately reporting the performance of your group to more senior levels of management.

8. Think about the alignment of these two groups of indicators and construct below the picture of the performance information you'd proceed with in two timeframes:
    - Now
    - In the medium term (12-24 months out).

9. For the performance information you'd proceed with now, try representing it in final form. For example, what would the information look like - graphic, narrative, diagrammatic, etc. - and how would you package it as management information?
   Questions of some use here include:
    - Who is the audience - internal management, clients/external scrutineers?
    - How would each audience like to receive the performance information?
    - How could you present it to accelerate understanding and to prompt worthwhile decision-making by your manager(s) or other audiences?
    - How often would you put forward the information and where would you place it in the management cycle?
    - What expectations would you have in terms of decision feedback and revision of the management information sent ?

10. You've now mapped out your best understanding of the information about performance worth moving through an information system to your managers, clients and external scrutineers.

You can use the following space to document:

a) Those performance indicators that you and your workgroup could use to correct or improve your own performance within your unit.
b) The information collection, analysis and interpretation changes you would need to plan for and implement if you are to put your medium term plans for performance information in place (refer to Question 7).

11. The work you've done to date in this worksheet has prepared your thoughts about performance information in a relative vacuum - in other words, it assumes you have managerial support and an enthusiastic workgroup.

... you are now ready to tackle the non-rational or political issues to do with performance information.

12. Makes notes about the interest in accountability and performance improvement your managers either:
    - Express, or
    - Demonstrate through their behaviour.

13. Assess the level of difficulty you might have in addressing performance information and improvement with:
    - Your managers
    - Your work unit

    In each case, look for the major impediments and what you think keeps those impediments alive or sustained:

    a) Your managers

Difficulty	Impediments	Sustained By

    b) Your Work Unit

Difficulty	Impediments	Sustained By

14. Taking these in turn, rough out the elements of strategy you could envisage that might make a sufficient shift in their level of interest and participation in performance improvement.
    Take from Question 12 those elements you'd need to target and where possible, specify what you might usefully attempt in regard to each:

    a) Manager

Target	Strategy

    b) Work Unit

Target	Strategy

15. From where else/whom could you hope to gain support:
    - Politically
    - Technically about
        - Performance Indicators
        - Management Information Systems

16. Draw together the threads of your notes here and briefly summarise:
    a) The context within which you believe you operate in regard to performance improvement and information.
    b) The approach you might take to having an impact on that context.
    c) The detailed action plans you'd need to develop if you were to proceed in your workplace.
    d) The pragmatic issues you will need to accommodate.

17. Show your summary to some people from a related work area and actively seek their constructive critique. Note their advice below.

Source: Program and Strategic Review Branch, `Guidelines for the Development of Performance Indicators', QLD Treasury, September, 1990.
(Note: This Worksheet has been condensed for economy in presentation and adapted slightly by the authors.)

FIGURE 6.8 Level of Operation: National. Sample performance indicators

Operational Function: Policy and Planning

Objectives	Indicators
• To provide literacy training to improve the level of skills of the community and employment opportunities	• Number of participants who enter employment or further education and training x months after completion of course: total participants completing course • Numbers of participants achieving entry level ALBE requirements for employment/vocational training/ further education: total number of participants • Population making literacy gains against national competence standards: total population • Rate of increase in program provision. • Numbers participating in programs: numbers in total population requiring assistance. • Number of participants: total population • Number of participants making literacy gains against national competence standards: total program participants • Number of priority group participants: total priority group • Total priority group participants: total participants • Australian workforce making literacy gains: total workforce
• To raise awareness of adult literacy needs and programs	• Number of states and systems including adult literacy goals and strategies in policies and strategies, including community development plans and economic and social policy: total states and systems • Number state policies and strategies including ALBE: total number state polices/strategies • Number of industry training plans which include literacy goals, issues and competencies: total industry training plans • Commonwealth portfolios and national systems identifying literacy in policies and plans: total commonwealth portfolios and national systems

FIGURE 6.8 (cont.) Level of Operation: National. Sample performance indicators

Operational Function: Policy and Planning

Objectives	Indicators
• To negotiate agreement on national benchmarks and strategies, including:- • industrial benchmarks - professional standards - curriculum - professional development	• Number states/territories/systems adopting national benchmarks: total number states/territories/systems • ABE practitioner competency standards • Tafe National Professional Development Strategy • Curriculum practice guidelines including: - good practice - inclusive practice • National Aboriginal and Torres Strait Islander Education Policy
• To undertake national research projects to improve understanding of adult literacy and to improve the quality of support for adult literacy	• National ALBE budget allocated to research: total national ALBE budget • Number nationally funded curriculum research projects tied to priority areas: total nationally funded curriculum research projects • Number nationally funded professional development research projects tied to priority areas: total nationally funded professional development research projects • Number states/territories/systems implementing findings of nationally funded research projects: total states and territories • Number of states/territories/systems incorporating research in strategic plans: total number states and territories

FIGURE 6.9: Level of Operation: National. Sample performance indicators

Operational Function: Program Development and Implementation

Objectives	Indicators
• To ensure that programs and services are consistent with national policies and priorities.	• Number of states and territories translating national policy into State education policies and strategies and implemented at provider level: number of states and territories • Number of states and territories developing support structures, including planning and co-ordination to avoid duplication: total number of states/territories • Number of states and systems developing structures for co-ordination and implementation of programs: total number of states and systems • Number of states/territories achieving national targets: total states and territories • Number of language and literacy programs linked with job retention, retraining and micro-economic reform: total number of programs • Numbers of places/enrolments/programs/program hours/successful outcomes in - Vocational Education - Industry Training - Adult community and further education : total number of places/enrolments/programs/program hours/successful outcomes
• Ensure priority groups have greater access to programs.	• Number of priority groups achieving goals: total achieving goals • Number of priority group enrolments: total enrolments • Number of priority group members participating in targeted programs, e.g. workplace: total participation in targeted programs • Number of program hours for priority groups: total number of program hours • **Number of participants receiving social and economic support: total number of participants.** • Number of states and territories developing consultative processes involving stakeholders: total number of states and territories
• Ensure removal of barriers to participation.	• Infrastructure funds allocated for childcare: total infrastructure funds • Infrastructure funds allocated for facilitated learning: total infrastructure funds • Infrastructure funds allocated for flexible delivery: total infrastructure funds

FIGURE 6.10: Level of Operation: National. Sample performance indicators

Operational Function: Personnel Management and Professional Development

Objectives	Indicators
• Development of national approach to professional development for ALBE personnel, including:- - development of national professional competency standards for ALBE practitioners - negotiation of nationally agreed industrial benchmarks for employment - negotiation of nationally agreed professional development strategy	• Number of States and systems implementing professional development policies and plans consistent with national priorities: total states/ systems • Number of States and systems committing funds to professional development: total states/systems • States and territories developing pre-service and in-service articulated and accredited professional development programs, consistent with national competency standards: total states/systems • Number of ALBE personnel achieving national competence standards: total ALBE personnel • Number of enrolments in professional development: number of successful completions in professional development programs. • Professional development funds: total program funds • Professional development funds: total salary funds. • Professional development funds: EFT tutor • Cost professional development: successful student outcome. • Number of states and systems working within national professional development strategy: total states/systems • Number of states and systems applying national industrial benchmarks for employment of staff • Number of national institutions identified and linked into strategic plans and policies: total national institutions • Number of professional associations identified and linked into policies and strategic plans: total professional associations

FIGURE 6.11: Level of Operation: National. Sample performance indicators

Operational Function: Curriculum

Objectives	Indicators
• To improve quality relevance and flexibility of programs so Australia can better adapt to changing economic, industrial and social circumstances.	• ALBE budget allocated to curriculum: total ALBE budget • ALBE budget allocated to curriculum research and innovation: total ALBE curriculum budget • Number of states and territories using nationally agreed benchmarks for language and literacy reporting: total states/territories • Number of states and systems using curricula and programs meeting national accreditation requirements: number states and systems • Number of states and systems developing curricula for specific contexts, including industry contexts: number states and systems • Level of literacy skills adult population current year : past year • Level of literacy skills in priority groups in adult population: level of literacy skills of adult population • Number of participants achieving goals: total number of participants. • Number of participants in priority groups advising goals: total number of priority group participants • Number of states and territories developing curriculum policies, priorities and plans: total number of states/territories • Number of states and territories developing infra-structure to support curriculum development: total number of state/territories • Curriculum development costs: total program costs • Cost of curriculum development: successful student outcome. • States and territories developing curricula using nationally agreed guidelines e.g. inclusive practice: total states and territories • States/systems developing policies and plans consistent with national policies: total states/systems • States/systems providing infrastructure support consistent with national policies: total states/systems • States/territories developing accreditation structures and procedures consistent with national guidelines: total states/territories

FIGURE 6.12: Level of Operation: National. Sample performance indicators

Operational Function: Information Services - Financial and Resource Management

Objectives	Indicators
- Efficient distribution of funds to States and Territories.	- Students x program hours: program cost - Successful outcomes: total program cost - Cost of social/economic assistance: total students (e.g number of FTA) - Cost of social/economic assistance: total students receiving assistance - Cost of social/economic assistance: successful student outcome. - Cost of social/economic assistance: program cost. - Number of correctly processed social/economic support requests e.g. FTA: total number of support requests.
- Development of a national information system which is flexible, reliable, efficient and secure and reflects strategic and policy goals.	- Number of states and systems providing accurate data as requested: total states/systems - Data elements reflecting national policy objectives and priorities: total data elements - Cost of data collection: number data elements collected

368 369

FIGURE 6.13: Level of operation: State/Territory. Sample performance indicators

Operational Function: Policy and Planning

Objective	Indicator
• To develop state policy for ALBE	• State education and training budget allocated for ALBE: total state education and training budget • Number national priorities and goals in state policy and plan: total national priorities and goals
• To promote public awareness of ALBE issues and provision	• Number of state industry training plans including ALBE needs and issues: total state industry training plans • Number of state departments and agencies adopting plain english documentation: total number state departments and agencies
• To increase participation	• Number of enrolments: total adult population requiring assistance • Number of enrolments: total adult population with language and literacy needs • Number of employers releasing staff for literacy/numeracy training: total number employers with staff requiring literacy/numeracy training
• To develop and expand program provision in priority and targeted areas - rural or isolated areas - under resourced areas - unemployed - NESB - Kooris - women - disabled - aged - youth	• Number of program hours in priority and targeted areas: total number of program hours • Number of student contact hours in targeted/strategic provision: total student contact hours • Enrolments in strategic and targeted provision: total enrolments • Number of priority group enrolments: total priority groups in population.

FIGURE 6.14: **Level of operation: State/Territory. Sample performance indicators**

Operational Function: Program Implementation and Development

Objective	Indicator
• To implement state and national policy	• Number of national goals and targets achieved e.g. Newstart enrolments achievement of ABE teacher competencies women: total national goals and targets • Number of state departments and agencies including adult literacy in policies and plans: total number of state departments and agencies
• To ensure access to adult literacy and basic education - childcare - flexible delivery - facilitated learning	• Infrastructure funds allocated to childcare: total infrastructure funds • Number funded childcare hours:number of program hours • Number funded childcare hours: number of childcare hours required • Number of funded facilitated hours:number of program hours • Number flexible delivery hours:total program hours • Flexible delivery enrolments:total enrolments • State education/training budget allocated to ALBE: total state education/training budget • ALBE budget allocated to policy/ management/infrastructure: total ALBE budget • Number of ALBE teaching officers: total ALBE non-teaching officers
• To provide appropriate levels of infrastructure and support • inclusion of ALBE in education/training structures - establishment of ALBE policy and management structures - recruitment and employment of ALBE professional officers • To provide accreditation structures and processes - ensuring state and national recognition and portability of credentials - providing pathways and clear articulation	• ALBE curricula meeting accreditation requirements: total ALBE curricula • Number of accredited curricula meeting industry standards and requirements: total accredited curricula • Number accredited ALBE curricula: number non accredited curricula • Number graduating from accredited courses: number enrolling in accredited courses. • Number accredited course graduates current year: number accredited course graduates past year

FIGURE: 6.16: Level of operation: State/Territory. Sample performance indicators

Operational Function: Curriculum

Objective	Indicator
• Maintain and improve the quality of programs and services	• ALBE budget allocated to curriculum: total ALBE budget • ALBE budget allocated to curriculum research and innovation: total ALBE budget
• To develop curricula consistent with state guidelines and nationally agreed benchmarks and principles	• ALBE curricula consistent with state curriculum development guidelines: total ALBE curricula • ALBE curricula meeting accreditation requirements: total ALBE curricula • Curricula consistent with national adult English Language, Literacy and Numeracy competency frameworks: total curricula • Curricula consistent with national good practice guidelines: total curricula
• Development and implementation of curricula to meet a range of learner goals and needs	• Number students demonstrating progress against national benchmarks: total number of enrolments • Students achieving employment,training and or personal goals:total number of students • Students achieving entry level qualifications (employment, vocational training, further education):total enrolment • Number student contact hours: total students • Number of priority group enrolments achieving goals: total priority group enrolments
• Development and consistent implementation of assessment tools and procedures	• Providers using agreed assessment tool: total providers • Providers participating in moderation and validation processes: total providers • Practitioners participating in moderation and validation processes: total practitioners • Number of paid moderation hours total EFT practitioners • Number of moderation hours: total per EFT practitioner

FIGURE 6.15: **Level of operation: State/Territory. Sample performance indicators**

Operational Function: Personnel Management and Professional Development

Objectives	Indicator
• Maintain and improve the quality of programs and services	• ALBE budget committed to Professional development: total ALBE budget • ALBE budget committed to paid participation in professional development: total ALBE budget
• To provide career pathways with in the ALBE field- including - range of teaching and non-teaching roles - consistent, equitable and secure conditions of employment	• Number of staff obtaining promotion: total staff • Number permanent/contract staff: number of casual sessional staff • Average length of employment • Staff receiving award/benchmark salaries and conditions: total staff • staff awarded study leave: total staff
• To provide and co-ordinate professional development for the ALBE field - - pre service - in service - post service - credentialled and - non credentialled	• Professional development programs meeting national/state competency benchmarks: total professional development programs • Number of staff with ALBE quals: total ALBE staff • Number of ALBE staff achieving national professional competency standards: total ALBE staff • Practitioners undertaking professional development: total practitioners • Number paid hours participation in professional development: total hours participation • Practitioners participating in paid professional development: total practitioners • Number of hours participation in professional development : total EFT practitioners • Number of hours paid participation EFT practitioners • Supply of qualified ALBE staff: total demand for qualified ALBE staff

FIGURE 6.17: Level of operation: State/Territory. Sample performance indicators

Operational Function: Information Services and Financial/Resource Management

Objectives	Indicators
• To ensure effective, efficient and equitable distribution of state and commonwealth funds	• Students x program hours: total cost of program • successful completion: total cost of program • Successful target group completion: total cost of program • Target group students x program hours: total cost of program • Total cost of professional development: total practitioners • Total cost of curriculum development: total students x program hours • State government funding : total funding • Non government funding : total budget
• To promote partnerships, resource sharing agreements and sponsorships	• Number of sponsorship/resource sharing agreements negotiated. • $value of resource sharing agreements: total budget. • Value of unpaid hours: total program operational costs • Value of volunteer teaching hours: cost of paid teaching hours
• To collect, maintain and aggregate data for commonwealth and state reporting purposes	• Number data collection items supplied to commonwealth and other agencies: total number data collection items requested • Number accurate data items collected: total data items collected • Number regions/providers/agencies contributing to data base: total regions/ providers/agencies

FIGURE 6.18: Level of Operation: Region. Sample performance indicators

Operational Function: Policy and Planning

Objectives	Indicators
• To implement national and state policies	• Number national priorities included in regional policy: total national priorities • Number state priorities included in regional policy: total state priorities • Number of national priorities achieved: total national priorities - target group enrolments - total enrolments - program areas and types • Number of state priorities achieved: total state priorities achieved • Funding tied to national targets: total funding • Funding tied to state targets: total funding • Providers including national/state/regional priorities in policy, plan, programs: total providers
• To identify and plan for priority areas: - Women - NESB - Kooris - Unemployed - Youth - Ageing - Disabled - Isolated - Rural fringe - Curriculum development - Professional development	• Establishment of regional consultative mechanisms and processes • Number of national priority groups in region included in consultative processes: total national priority groups in region • Number of state priority groups in region included in consultative processes: number of state priority groups in region • Number of priority areas tied to profile of region: number of priority areas • Number of priority group enrolments: total number of priority group in population. • Current year enrolments: past year enrolments • Current year student contact hours: past year student contact hours • Current year hours of provision: past year hours of provision • Number of new enrolments: total enrolments
• To provide and expand adult literacy and basic education in the region	• Number of enrolments: regional adult population • Number of enrolments: regional adult population requiring assistance • Number of new programs/program hours: total programs/program hours
• To promote community awareness	

FIGURE 6.19: Level of Operation: Region. Sample performance indicators

Operational Function: **Program Development and Implementation**

Objective	Indicators
• To review and plan program provision to meet regional needs	• Number of providers meeting national, state, regional objectives : total providers • Number of providers targeting national, state, regional objectives: total providers • Funds allocated for seeding/innovative programs: total funds • Budget allocated for priority groups: total budget
• To offer advice and educational leadership to providers	• Funds allocated for infrastructure support: total funds

FIGURE 6.20: Level of Operation: Region. Sample performance indicators

Operational Function: Personnel Management and Professional Development

Objective	Indicators
• To provide career pathways within the ALBE field including: - range of teaching and non teaching roles - consistent, equitable and secure conditions of employment	• Positions funded according to award/ benchmark salaries and conditions: total positions funded • Funded EFT non-teaching positions: total EFT non-teaching positions • Funded EFT teaching positions: total EFT teaching positions • Number permanent positions:total number of positions • Number funded EFT teaching staff: number funded EFT non-teaching staff
•To provide and co-ordinate professional development	• Number of providers working to professional development policy and plan: total number of providers • Regional allocation for paid participation in professional development: total regional budget • Regional budget allocated for professional development : total regional budget • Regionally funded professional development programs meeting national/state/regional priorities: total regionally funded professional development • Professional development programs meeting national/state competency benchmarks: total professional development programs • Practitioners participating in regional professional development: total practitioners • Number of hours paid participation in professional development: total EFT practitioners • Number of hours participation in professional development: total EFT practitioners • Number of providers allocating funds to professional development: total providers • Average provider budget allocated to professional development: total provider budget

FIGURE 6.21: Level of Operation: Region. Sample performance indicators

Operational Function: Curriculum

Objective	Indicators
• To improve quality of curriculum program and services	• Regional budget allocated to curriculum: total regional budget • Providers allocating funds to curriculum development: total of providers • Average provider budget allocated to curriculum development: average total provider budget • Number curriculum documents meeting national/state benchmarks: total curriculum documents - accreditation requirements - good practice guidelines:
	• Budget allocated to curriculum and assessment moderation and validation: total budget • Number hours provision of paid moderation and validation: total effective full time practitioners • Paid moderation and validation hours used: total paid moderation hours available • Providers using moderated assessment processes: total providers • Providers using recognition of prior learning assessments: total providers
• To identify curriculum priorities and gaps	• Regional budget allocated to curriculum development and innovation: total regional budget • Regional budget allocated to resources and materials: total regional budget • Number of resources produced, which meet national/state benchmarks: total resources produced • good practice guidelines • priority areas • Number of curriculum documents produced tied to priority areas: total number curriculum documents • Number curriculum documents meeting state and national guidelines: total curriculum documents • Number of providers using accredited curriculum consistent with state and national guidelines: total providers • Number of providers developing internal learner pathways: total number of providers • Number of providers with programs clearly articulating with programs in other agencies and total number of providers.

FIGURE 6.22: Level of Operation: Region. Sample performance indicators

Operational Function: Information Services and Financial/Resources Management

Objective	Indicators
• To collect and maintain accurate statistical information for state and commonwealth agencies	• Number accurately collected data elements • Ratio of providers and agencies contributing to regional data base: total agencies and providers • Number of data items provided to state and commonwealth: data items required by state and commonwealth
• To monitor the efficient distribution and use of regional funds	• Number of providers meeting requirements of funding and service agreements: total providers • Total program hours x total enrolments: expenditure • Number successful outcomes: total costs • Funds allocated to national priorities: total expenditure • Funds allocated to state priorities: total expenditure • Total successful national priority outcome: total expenditure • Total successful state priority outcome: total expenditure • Total professional development expenditure: total successful outcomes • Total curriculum development costs: total student contact hours • Total cost of curriculum development: total successful outcomes • Total cost of student support: Total successful outcomes
• To widen the funding base including: - partnerships between levels of government - partnerships between providers, community agencies - partnerships within industry - cost recovery and fee for service agreements	• Total state funded regional expenditure on adult literacy and basic education: total regional expenditure on adult literacy and basic education • Cost recovery funding/fee for service income for ALBE: total regional income for ALBE • Number of resource sharing agreements between providers, community agencies • Dollar value of resource sharing agreements between providers

FIGURE 6.22 (cont.): **Level of Operations: Region. Sample performance indicators**

Operational Function: Information Services and Financial/Resources Management

Objective	Indicators
• To use state and commonwealth funds for maximum impact	• Regional expenditure on infrastructure: regional expenditure on programs • State funded regional expenditure on infrastructure: commonwealth funded regional expenditure on infrastructure • Commonwealth funded regional expenditure on infrastructure: commonwealth funded regional expenditure on programs. • State funded regional expenditure on infrastructure: Commonwealth funded expenditure on programs • Fee for service regional expenditure on infrastructure: fee for service regional expenditure on programs

390

391

FIGURE 6.23: Level of Operation: Provider. Sample performance indicators

Operational Function: Policy and Planning

Objectives	Indicators
• To implement state and national policies	• Number of national and/or state policy priorities identified in provider plan and program: total number of national/state priorities
• To improve the participation rates of targeted social justice and equity groups	• Participants each target group: total adult target group • Target group participation: total participation • Target group enrolment: total enrolments • Number of staff, and members of advisory groups and management committees in target groups: total number of staff, members of advisory groups and management committees • Number of targeted programs: total programs • Enrolments in targeted programs: total enrolments • Student contact hours achieved by target groups: total student contact hours
• To increase awareness of education and training staff about literacy needs of adults	• Number of college departments making referrals or participating in programs • College budget funds allocated to ALBE programs: total college budget • Number of referrals from college departments: total referrals
• To increase community awareness of literacy needs of adults	• Number of responses to promotional material • Number/type of community organisations, employer groups, unions, making referrals to programs • Number/type of community organisations, employer groups, unions liaising with literacy provider • Number of employers including literacy in staff training: Total local employers.
• To promote and market provision and raise community awareness of its availability	• Program participants: total adult residents in local government area
• To liaise and network in the promotion of adult literacy	• Number of staff participating in local, regional, statewide and national ALBE activities for networking and exchange: total number of staff

FIGURE 6.24: Level of Operation: Provider. Sample performance indicators

Operational Function: Program Development and Implementation

Objectives	Indicators
• To provide programs to meet local needs and objectives	• Number of programs planned/offered with other community, TAFE and industry providers to meet local gaps in provision: total programs • Number of satisfied industry clients: total industry clients.
• To identify gaps and unmet needs	• Number of targeted program hours tied to local need : total number of program hours • Number of programs introduced to absorb waiting lists: total programs • Number of programs offered in new areas: total programs: - workplace - vocational/literacy mix - literacy support
• To develop appropriate modes of delivery	• Number of distance programs: total programs • Number of open learning programs: total programs • Number of part-time programs: total programs • Number of full-time programs: total programs • Number of small group programs: total programs • Number of one to one programs: total programs • Number of workfplace/industry programs: total programs

FIGURE 6.25: Level of Operation: Provider. Sample performance indicators

Operational Function: Personnel Management and Professional Development

Objectives	Indicators
• To facilitate professional development and in-service of staff - development of professional development policy and plan	• Development of professional development policy and plan • Budget allocation for professional development: total budget allocation • Practitioners participating in provider, regional, state, professional development: total practitioners • Support for professional development - Number of childcare hours: number professional development hours - Number of sessions attracting travel funds for participants: total number of sessions - Number paid practitioner participation hours: total practitioner participation hours in professional development • Number of professional development sessions linked with state/national priorities: total professional development • Number of professional development hours: number EFT practitioners • Number of readable and useable memos, bulletins, newsletters, produced/distributed: total produced • Number of memos, bulletins and newsletters used: total produced • Number/accessibility of network meetings • Number of practitioners adopting/using new methods and resources: total number of practitioners

FIGURE 6.25 (cont.): Level of Operation: Provider.Sample performance indicators

Operational Function: Personnel Management and Professional Development

Objectives	Indicators
• To provide co-ordination and support for volunteer tutors	• Number of staff resignations per year: total number of staff employed per year • Rate of staff attrition • EFT teaching staff: EFT non-teaching staff • Permanent teaching staff: non-permanent/sessional teaching staff • F.T. teaching staff: part-time/sessional teaching staff • Number of staff expressing satisfaction with workplace: total number of staff • Staff expressing satisfaction with leadership of management/co-ordinator: total staff • Staff identifying own professional development needs: total staff • Staff undertaking additional professional study and training: total staff • Staff applying for and/or gaining career advancement: total staff • Teaching staff achieving professional standards (national ABE teachers competency standards, national Workplace ABE teacher competency standards): total teaching staff • Number of staff receiving award/benchmark salaries and conditions: total staff
• To provide career paths and work satisfaction for staff	• Retention rates of volunteer tutors • Recruitment rate of volunteer tutors • Number of new volunteers: number continuing volunteers • Number of tutors offering extra hours : total volunteers • Number of volunteers participating in professional development : total number of volunteers • Hours of volunteer training: number of volunteers • Funds allocated for reimbursement of volunteers out of pocket expenses: total budget • Number of hours provided by volunteers: total program operation hours • Number of tuition hours provided by volunteers: total number of program delivery hours

FIGURE 6.26: Level of Operation: Provider. Sample performance indicators

Operational Function: Curriculum

Objectives	Indicators
• To provide high quality programs for learners	• Budget allocated for curriculum development: total budget • Budget allocated for professional development: total budget • Number of practitioners participating in professional development: total practitioners • Number of graduations/completions of professional development programs: enrolments in professional development programs • Number of students completing programs: total number of students • Number of students achieving goals: total number of students • Number of programs with clear articulation and pathways: total programs • Cost of resource collection: total students • Cost of resource collection: total practitioners • Number of student contact hours: total number of students.
• To ensure the effectiveness of the program in meeting student aims by attending to: - curriculum including assessment issues - monitoring progress and evaluating against program aims	• Number of assessments made using recognised, validated tool: total number of assessments • Number of hours of paid assessment moderation and validation training: total EFT practitioners. • Number of hours of paid moderation: number of courses. • Number of practitioners participating in moderation: total number practitioners
• To clearly identify and advertise selection and placement criteria	• Number of courses with published entry criteria : total courses
• To clearly advertise program/course objectives/outcomes and style of provision	• Number of students choosing courses related to program/course objectives: total students • Number of students enrolling because of course objective: those enrolling because of convenience of timetabling, venue. • Number of course outlines/objectives published: total number of courses
• To meet individual student goals	• Number of students making language, literacy and numeracy gains against state/national benchmarks: total students

FIGURE 6.26 (cont.): Level of Operation: Provider. Sample performance indicators

Operational Function: Curriculum

Objectives	Indicators
• To attain class goals and outcomes	• Number of students graduating from accredited programs: total enrolling in accredited programs. • Number of students expressing satisfaction with course: total students • Number of students achieving goals: total students • Number of students achieving employment goals: total students • Number of students achieving further education goals: total students • Number of students achieving vocational training goals: total students • Number of students achieving personal goals: total students • Number of students achieving learning goals: total students • Number of students completing particular course/module/subject: total students enrolling in course/module/subject • Number of students completing particular course/module/subject: total students enrolling in course/module/subject.
• To clearly identify and state student assessment processes	• Use of common reporting terminology with accreditation and competency frameworks • Number of students undergoing entry, on course and exit interviews : total number students • Number of students using self-evaluation procedures : total number of students
• To keep accurately updated student learning needs and progress records to assist program planning and review of provision and point to areas requiring new initiatives.	• **Time allocated to maintenance of maintenance of entry, on course and exit interview data: total administration time**
• To provide educational counselling and referral services	• Hours of educational counselling: client contacts • Hours of educational counselling: total students • Hours of educational counselling: total student successful outcomes
• To clearly state proposed course outcomes	• Number of programs with clear entry and exit criteria: total programs • Number programs articulating into: - ABE - Further education - Vocational training - Entry level requirements for/or employment

FIGURE 6.26 (cont.): Level of Operation: Provider. Sample performance indicators

Operational Function: Curriculum

Objectives	Indicators
• To provide literacy/numeracy provision as negotiated with clients - program type - delivery - program goals	• Number of programs offered in response to student requests : total programs offered • Number of programs with clear articulation to other internal or external programs: total programs. • Number of satisfied industry clients: total industry clients • Number of industry clients reporting achievement of specific goals: total number of industry clients. • Number of industry clients reporting positive unplanned outcomes: total number of industry clients. • Number satisfied DEET agencies contracting programs: number of DEET agencies contacting programs • Number of satisfied students: total students
• To provide tuition via flexible approaches and/or time tables and venues	• Number of venues and locations for localised need: venues/locations. • Number of program operating hours operating on day/eveningshifts basis: to total program hours.
• To develop appropriate and relevant curriculum for particular student groups, with attention to: - Content - Gender/ethnic inclusivity - Mode of delivery - Assessment criteria - Flexible delivery for distance, isolated and incarcerated groups	• Number of curricula meeting national or state standards : total curricula • Number of accredited courses offered : total courses offered • Number of curriculum documents using state/national guidelines : total curriculum documents • Number of curriculum documents: number of programs offered in provider • Number of curriculum developed for specific learning goals proposed: total curriculum

FIGURE 6.27: Level of Operation: Provider. Sample performance indicators

Operational Function: Information Services and Financial/Resource Management

Objectives	Indicators
• To achieve efficient use of financial and physical resources	• Cost per student contact hour • Cost per course completion/graduation: total students • Student contact hours: teaching hours • Hours room usage: hours room available • Increase in funding: increase student contact hours • Cost of resource collection: total students • Cost of resource collection: total practitioners • Hours of use of physical facilities and resources: hours available • Hours of unpaid work: total hours of work
• To maintain accurate financial records	• Budgeted expenditure: Actual expenditure • Cost recovery income: total income • Tendered program income: total income • Rate of increase in program funding • Total income current year: total income past year
• To maintain accurate student data records	• Time allocated for maintaining records: total hours employed • Key data elements collected at enrolment: total key data elements required
• To increase shared use of physical resources and staff between colleges, community providers, other providers and community agencies.	

FIGURE 6.28: IMPACT EVALUATION - QUESTIONS TO ASK

	QUESTIONS TO ASK	FOCUS OF EVALUATION	KEY AUDIENCES FOR EVALUATION
NATIONAL	Were national goals and targets achieved? Is the expenditure justified in terms of gains in literacy and national productivity? Was this the result of the policy initiatives or were other factors involved? How was policy implemented? Have some States been more successful in implementing their program than others? Are national literacy needs being met? Is the policy still appropriate? Are the goals still appropriate?	Existing Programs	Federal and State Ministers and portfolios.
STATE/TERRITORY	To what extent were the intended objectives achieved? How successful were programs in reaching the total target area/group? How does the current level of service compare with the need for the service? Were some programs more effective than others in achieving objectives? Did some programs have unanticipated consequences?	Existing Programs	Policymakers and Planners Funding bodies
REGIONAL	To what extent have regional goals been achieved? Are the real goals reflected in policy statements? Have some programs performed better than others in achieving desired outcomes? Should some programs be encouraged over and above others? Have there been any unanticipated outcomes, desirable or undesirable as a result of programs? What are the effects of professional development strategies on program performance? Is the present level of expenditure on professional development justified?	Existing Programs	Policymakers and Planners Funding bodies
PROVIDER	Are we clear about the real objectives of the program? Are these objectives being achieved? Are the students learning the full range of skills as set out in our curriculum policy? Are the skills gained achieved in sufficient depth to be useful in their future lives? What are the short-term and long-term outcomes for our students? Do different programs cater for some categories of students better than others? Are programs satisfying the full range of needs of our students?	Existing Programs	Funding agencies Program Managers and Staff

Chapter 7 INFORMATION COLLECTION ALTERNATIVES

The *overview framework* introduced in Chapter 2 (as Figure 2.2) identifies a range of information collection alternatives and highlights their advantages and disadvantages. These alternatives are categorised according to whether the information is gathered:

- direct from individuals
- by an independent observer
- using mechanical devices
- via unobtrusive techniques
- from existing records.

The categories are used for ease of discussion and are not mutually exclusive. For example:

- rating scales may be used in collecting information direct from individuals as the source or when information is being compiled by an external observer
- unobtrusive measures might include private records already in existence such as diaries written by individuals.

It is not possible to provide an in-depth and comprehensive understanding of each of these information collection alternatives here. For example, the design of questionnaires is an art in itself and entire books have been written on the subject. A list of references in specific topics has been included for those who wish to pursue them in more detail.

As indicated in the *overview framework* in Figure 2.3, the selection of information collection techniques depends on the:

- key evaluation issues
- availability of resources
- time-constraints related to the use of the findings.

It is important to know that collection of information from more than one source (using more than one method) enhances the confidence one can have in the 'truth' or validity of the findings. It is also important to consider the demands that information collection makes on those involved. Where information collection requires time and effort on the part of respondents, the quality of that information is directly related to their empathy with the project and their willingness to cooperate. The range of alternative data collection methods are discussed in the following sections.

1. Information Directly from Individuals

Information direct from individuals can be gathered via 'self-reports' or 'personal products'.

A. Self-Reports

a. Written responses, diaries or anecdotal records.

Written responses and diaries allow individuals to describe unique situations in their own words and to focus on those program issues or concerns which they believe are important. Anecdotal records are simply reports about incidents or aspects which are believed to reflect something typical or significant about the program.

Working with self-reports can be difficult in that they are often not well maintained, the quality of response depends on the skills and attitudes of the individuals involved, and the resulting data can be difficult to interpret.

If written records are required by the evaluation but are not part of normal program operations, they may involve additional work by program staff. This may be strongly resented and can pose a significant threat to the quality and validity of information collected.

b. Checklists or Inventories.

Checklists or inventories consist of a list of items and respondents are asked to tick or circle each relevant item. An advantage in using them is that they are quick and easy for respondents to complete and evaluators to score.

Unfortunately, there is no opportunity in this approach for individuals to qualify their answers. The response format is described as 'closed' because respondents are forced to make a choice among the items provided (see further discussion on open and closed formats later in this chapter). Furthermore, the design of the checklist can lead to false opinions either by giving respondents too few alternatives or by prompting socially acceptable answers. It is necessary to put a lot of thought into developing alternative responses. 'This can be done by careful pilot testing using less structured approaches to locate the range of likely responses and by using the category called 'other (please specify)' to allow for unanticipated responses.' (de Vaus, 1990: 86)

c. Rating Scales

A rating scale provides respondents with statements and asks them to indicate how strongly they agree or disagree with them (Likert-style format). As indicated in Figure 7.1, the format for this may be verbal or diagrammatic. The rating scale has similar advantages and disadvantages to the forced-choice format of the checklist or inventory approach described above.

The rating scale is frequently used in a strategy described as the 'Delphi Technique'. This has been defined as 'a method of systematically, soliciting, collecting, evaluating and tabulating expert opinions, usually in long-range forecasting' (Morrison, in Stanley: 1987). The Delphi technique involves:

- collating the written ideas of experts on issues of current or potential importance
- mailing the collated ideas to each of the experts involved, asking them to assess their relative importance and inviting them to add to them
- identifying areas of agreement and disagreement, and repeating the process, asking respondents to reconsider their answers and revise them if necessary on the basis of their reactions to the responses of their colleagues
- repeating the process until a consensus is achieved (Dixon & Harding, 1990:256).

The Delphi technique:

- is conducted anonymously so that the focus is on the ideas. Individual respondents' opinions are not influenced by powerful advocates, and they can alter their opinions in the light of the views of others without 'losing face'.

- can draw on the opinions of large numbers of people without the need for them to be available for meetings.

While these can be seen as advantages, the Delphi technique is also time-consuming, taking up to one to three months to complete. This represents a substantial amount of work for those involved in mailing questionnaires, analysing data, and sustaining the support of those involved over the period. Figure 7.2 provides a set of guidelines for conducting a Delphi study.

Further Reading:

Curriculum Branch, Education Department of Victoria (1986) Destination: Decisions, Decision-making Strategies for School Communities, Education Department of Victoria.

Cyphert, F.R. & Gant, W.L. (1971) 'The Delphi Technique: A Case Study', Phi Delta Kappan, 272-3.

Delbecq, A.L., Van De Ven, A.H. & Gustafson, D.H. (1975) Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes, Scott, Foresman & Company, USA.

Dixon, J. & Harding, G. (1990) 'Delphi Forecasting Methodology', Workshop Paper, Conference Proceedings, National Evaluation Conference.

Linstone, H.A. & Turnoff, M. (eds) (1975) The Delphi Method: Techniques and Applications, Addison-Wesley, Massachusetts.

d. Ranking Scales

Ranking involves making a list of things such as ideas or activities and asking respondents to place them in order from highest to lowest, most positive to most negative and so on. Ranking does not reflect the extent of the differences between the items. Figure 7.3 provides an example of a ranking scale in which respondents are asked to rank a set of items. Like rating scales, they represent a closed or forced-choice format. Rankings can be unreliable, particularly the middle items, if the list of items is long (de Vaus, 1990:88).

A Q-sort involves respondents in ranking groups or clusters of items. Below is an example of a Q-sort involving a set of curriculum goal statements:

> "In this case, you are asked to select what you see as being the most and least important goals of the curriculum. Because of their significance, probably none of these goals will be totally unimportant to you. Your task is to take the thirteen goal cards and then place each in one of five piles labelled 'Most important', 'Important', 'Somewhat important', and 'Least Important'. You should choose only one goal as most important and one as least important. You may choose three goals as very important, three as somewhat important and five as important. Make certain you place all thirteen goal cards on the basis of your own personal preferences." (Education Department of Victoria, 1985:94)

e. Semantic Differentials

Semantic differentials, like those included in Figure 7.4, are used to explore how people are feeling about an issue. Adjectives are chosen to represent the two extremes of a continuum and respondents are asked to indicate their feelings by placing a mark or selecting a number between the two extremes. If used properly they can be employed to determine a spectrum of attitudes towards key characteristics of a given program.

f. Questionnaires (Self-Administered)

Questionnaires are 'instruments that present information to a respondent in writing or through the use of pictures and then require a written response - a check, a circle, a word, a sentence, or several sentences.' (King, et.al., 1987:46) The questionnaire is a highly structured data collection technique in which each respondent is presented with a similar, if not the same, set of questions. Advantages of questionnaires are that they can be used to access a large number of people, they are cheap to administer and respondents can complete them at their own convenience.

However, questionnaires do not offer the flexibility of personal interviews, they require a level of skill in both reading and writing on the part of the respondent, and there is usually time and effort involved in following up those who do not respond.

Designing a questionnaire requires considerable expertise. When questionnaires are self-administered, as opposed to being administered by a trained interviewer, it is essential to concentrate on simplicity and clarity in both lay-out and question design.

Fowler (1984: 99 -106) suggests that there are four basic steps in the design of a good questionnaire:

- deciding what is to be measured. This involves clarifying the research objectives, defining precisely the variables that are to be measured or the phenomenon to be described and drafting an analysis plan.

- conducting focussed discussion with a small group of people representative of the population which is to be studied. This discussion is designed to provide insights into the issues or concerns related to the area that is to be investigated and provides the basis for developing specific questions.

- designing the questions, format and layout (see further discussion below).

- pilot-testing the questionnaire. A mailing to a sample of potential respondents may provide useful estimates of rate of return but will not necessarily give a full insight into difficulties in questionnaire design. A more effective method of pilot testing self-administered questionnaires is to organise a small group of potential respondents to complete the questionnaire with an evaluator present. This provides an indication of the length of time the questionnaire requires for completion. A follow-up discussion with respondents can then be used to identify areas of confusion, ambiguous questions and clarify the meaning of their responses.

Designing questions

When designing questions it is important to consider:

- content
- wording
- type i.e. whether it should be open or closed (de Vaus, 1990: 81-88).

(i) The Content

There are four types of question content: attributes, behaviour, beliefs and attitudes (Dillman, 1978:80). These are described below and illustrated by examples relevant to ALBE.

Attributes: These questions are designed to provide basic information about the respondent, information that will vary according to the purpose of the evaluation. However, such questions will often include information about age, occupation, gender, ethnicity and so on. For a literacy program evaluation, it might include information about the length of time participants had been in the program, the number of trainers assigned to them, and how much time per week they were involved with working directly with their trainers, etc.

Behaviour: Behavioural questions attempt to find out what the respondent actually does. For example, 'Who do you talk to if you have problems with your work ?' However, it should be noted that behaviour patterns are not always an accurate reflection of belief systems.

Beliefs: Belief questions attempt to discover what people think is true or false. For example, 'Do you believe that numeracy skills are important for your job?'

Attitudes: Attitude questions aim to establish what people think is desirable. For example, 'Do you feel that there should be more opportunities for working with others in small groups?'

It is important to distinguish clearly between the four different question types for a number of reasons:

- it ensures that questions are tapping information that is consistent with the purpose of the evaluation
- all four types of questions might need to be included to ensure that the topic being studied is systematically examined
- it is inappropriate to combine the different question types because they measure different things, whereas, for example, attitude questions can be combined with other attitude questions to develop a composite measure (see de Vaus, 1990:81-83).

(ii) Wording

De Vaus (1990:83-85) provides the following checklist of 14 questions designed to highlight the most obvious problems with question wording:

1. Is the language simple?
2. Can the question be shortened?
3. Is the question double-barrelled? (i.e. does it ask more than one question at once?)
4. Is the question leading? (i.e. respondents should be able to respond freely to questions without, for example, being influenced by the use of loaded words or only a limited range of alternative answers.)
5. Is the question negative? (i.e. questions using 'not' can be confusing.)
6. Is the respondent likely to have the necessary knowledge?
7. Will the words have the same meaning for all respondents?
8. Is there a prestige bias in the question (i.e. will respondents be inclined to give 'socially desirable' responses?)
9. Is the question ambiguous?
10. Do you need a direct or indirect question ? (i.e. on some sensitive issues it may be offensive to ask a direct question.)
11. Is the frame of reference for the question sufficiently clear? (for example, if you ask how often a student works on his/her own, you will need to give a time-frame such as 'per day' or 'per week'.)
12. Does the question artificially create opinions? (i.e. respondents should have the option not to give an opinion. Forcing them to select from alternatives which

do not reflect their beliefs will invalidate the responses.)

13. Is personal or impersonal wording preferable? (i.e. do you want to ascertain what the individual feels about something, or their perception of other people's attitudes.)
14. Is the question wording unnecessarily detailed or objectionable? (for example, asking respondents to place themselves in age or income categories may meet with less resistance than asking them their precise age or income).

(iii) Question Type

Decisions must be made about response format i.e. whether questions should be 'closed' or 'open'. Closed or forced choice questions have advantages and disadvantages and have been discussed above in relation to 'rating scales', which are one form of the closed response format. Other closed response formats discussed above include checklists, semantic differentials and ranking scales. Open questions have the advantage of allowing free uninhibited responses, but are time-consuming to analyse and the opinions of the more articulate/fluent respondents may unduly influence the findings.

The choice of open versus closed questions will depend on a number of factors:

- question content
- respondent motivation
- method of administration
- type of respondents
- access to skilled coders to code the responses to open questions (de Vaus, 1990:87).

Figure 7.5 provides an overview of different response formats used in questionnaires. Figure 7.5a provides an example of an open-response questionnaire, Figure 7.5b is a questionnaire which relies solely on closed responses and Figure 7.5c combines the two response formats. Gallup (1947) proposes some guidelines for combining the two in exploring participant response to a specific issue:

- a closed question can be used to determine whether the respondent is familiar with the issue and/or has thought about it.
- an open question can be used to explore their general feelings.
- a closed question can be used to ascertain specific information about the issue.
- open or closed questions can be used to tap respondents' reasons for feeling the way they do.
- a closed question can be used to identify how strongly they hold their opinions.

During pilot testing the questionnaire, a number of checks should be made for each of the questions (de Vaus, 1990: 88-89):

- the ability of the item to discriminate i.e. if all respondents answer the same way to a question it may not be worth asking.

- if combined sets of items (or 'scales) are used, it is important they measure the one concept (see further discussion of scales in Chapter 8).

- each item needs to be examined for reliability which means that the same result is achieved on repeated occasions, and validity which means that the measures actually measure what they are intended to measure (see further discussion of reliability and validity in Chapter 8).

- to avoid redundant questions i.e. two questions or items measuring virtually the same thing. This is done by computing the correlations between the two items. De Vaus suggests that if the items both measure the same concept and correlate over 0.8, then one of the items can be excluded.

- to avoid respondents developing a 'response set'. For example, if respondents establish a pattern of agreeing with the initial questions, they may continue with this pattern. Where possible, questions should be ordered to avoid an 'acquiescent response set' developing.

Once the questions addressing the evaluation issues have been developed, the actual layout of the questionnaire needs to be planned. This involves attention to the following key areas:

- answering procedures
- contingency questions
- instructions
- use of space
- order of questions
- setting up for coding.

Figure 7.6 provides guidelines supplied by de Vaus (1990: 89-93) in each of these six areas.

Further Reading:

Bradburn, N.M. & Sudman, S.(1979) Improving Interview Method and Questionnaire Design, San Francisco, Jossey-Bass.

Oppenheim, A.N.(1968) Questionnaire Design and Attitude Measurement, London, Heinemann.

Payne, S.(1951) The Art of Asking Questions, Princeton: Princeton University Press.

Shaw, M.E. & Wright, J.M. (1967) Scales for the Measurement of Attitudes, McGraw-Hill, New York.

Schuman, H. & Presser, S. (1981) Questions and Answers in Attitude Surveys, New York, Academic Press.

Sudman & Bradburn (1982) Asking Questions: A Practical Guide to Questionnaire Design

B. Personal Products

Information can be collected about the knowledge and skills of individuals via written achievement tests and samples of work. A difficulty with written achievement tests is that they do not necessarily reflect the individual's ability to transfer knowledge and skills to other contexts.

Samples of work are usually readily available and provide an indication of ability to respond to a specific task in a specific context, although like written tests, they do not necessarily reflect the individual's ability to transfer and apply knowledge and skills.

Determining learning outcomes in the ALBE field may involve the assessment of specific criteria in the context of a curriculum by analysing samples of work.

The measurement of learning outcomes is a complex issue and the question of individual learner assessment is not dealt with here. It is, however, important to stress that in the context of evaluation, the assessment of learning outcomes is used to make statements about the effectiveness of a program designed to lead to the outcomes.

As we have indicated previously, a national project has been devoted to developing a framework specifically for describing competencies in adult English language, literacy and numeracy (ACTRAC:1993) which are central to competence for work and social activity. This framework is based on the premise that competence involves a connection between performance and knowledge and skills which permits the achievement of social goals in particular contexts. It stresses the importance of both 'knowing' and 'doing' and the use of multiple forms of assessment (ACTRAC, 1993:27).

2. Information Compiled by an Independent Observer

The information compiled by an independent observer might take one or more of the following forms:

- Written Accounts
- Observation Forms
- Oral Responses

A. Written Accounts

Some evaluations call for detailed written descriptions of program activities, processes and participants. This is done by going into the field to observe the program in its natural setting. The main strength of naturalistic observation is that data are being collected where the action is as it occurs (Patton, 1978:72). This has a number of advantages in that it:

- permits an understanding of the context within which the program activities occur

- allows the evaluator to experience the program directly rather than relying on participant descriptions of the program

- may highlight aspects of the program which might not be noticed by those actually participating in it

- may identify some sensitive issues which participants are reluctant to raise or discuss in an interview situation

- allows the evaluator to go beyond the selective perceptions held by program participants (see Patton, 1978:73).

In some evaluations, it is appropriate for the observer to take a naturalistic stance and go into the field without any preconceptions about the program. In other evaluations, the observer may well use the program plan to establish how well the program in action reflects that plan. Three features common to all programs might be used to develop an outline of the program and its critical characteristics:

- the context, or tangible features of the program and its setting (for example, the classrooms, the resources used, the number and types of students, etc.)

- the underlying program logic concerning intentions and effects. For example, a Workplace Basic Education program might be based on the implicit or explicit assumption that if employees are taught literacy skills in their specific work context their productivity will improve.

- the activities and how the program was implemented. This will involve answering questions such as:

- 'What materials were used? Were they used as indicated?
- What procedures were prescribed for program staff in their interactions with clients? Were these procedures followed?
- In what activities were program participants supposed to participate? Did they?
- What administrative arrangements did the program include? What lines of authority were to be used for making important decisions? What changes occurred in these arrangements?' (King, 1987:28-29)

B. Observation Forms

The purpose and focus of evaluation will influence the type of approach taken to observation in the field. These are:

- the extent to which the evaluator participates in the program
- the degree of awareness of participants about the presence and role of the evaluator
- the degree to which participants are informed about the purpose of evaluation
- the frequency and duration of the observations
- the degree to which the focus is narrowly defined or intended to encompass a large range of program activities.

Figure 7.7 provides an explanatory summary of these five dimensions (Patton, 1987:81).

While informal observation of program implementation is useful, it may not be credible as evidence to some stakeholders. In many instances, it is highly desirable that observation methods are systematically planned and tailored to the purposes of the evaluation. The nature of the evaluation and the program setting dictate the amount of structure that is required when making observations.

Three different types of 'observation instruments' are described here:

- on-the-spot checklists which are used for recording the presence or absence or frequency of certain program activities or types of participant behaviour.

- coded behaviour records which involve assigning a particular behaviour a symbol and marking down the symbol whenever that behaviour occurs. For example, 'A' might represent teacher, 'r' might represent requests and 'v' might represent volunteer. Each time the teacher requested the help of the volunteer, the observer would record Arv. This recording system allows for a detailed account of program activities to be developed using a system of codes.

- delayed reports which involve the observers completing questionnaires about the program immediately after the observation period is completed. Observers should be familiar with the questions they will be asked, as well as having a formal definition of what is to be observed in a given time-frame (King, 1987: 91-93).

It is sometimes possible to use an observation instrument which has already been developed if the evaluator can be satisfied that the instrument has 'face validity' (see Further Reading below).

Further Reading:

Australian Commitee for Training Curriculum (1993) National Framework of Adult English Language Literacy and Numeracy

Competence. ACTRAC Products Ltd., Frankston, Victoria.

Good, T.L. & Brophy, J.E. (1984) Looking in classrooms (3rd.ed) Harper & Rowe, New York.

Stallings, J.A. (1985) Stallings' observation instrument: Training Manual, Nashville, TN: Vanderbuilt University, Peabody Center for Effective Teaching.

C. Oral Responses (either singly or in groups)

Oral responses can be collected via:
- face-to-face interviews
- telephone interviews
- group interactions (for example, nominal group forums and search conferences).

a. Face-to-face Interviews

Interviewing involves asking questions, recording the answers and possibly probing for additional responses. 'The persons interviewed may be program participants, staff, administrators, community members, funders or officials. In each case, the evaluator is seeking to find out how that person views the program under study.'(Patton, 1987:109) Interviews help to explain outward behaviour by attempting to enter into the other person's perspective to explore their thoughts, feelings and intentions.

Interview formats vary according to the extent to which they are structured before the interview is conducted. Patton (1987:116) identifies four different types: informal conversational interview, interview guide approach, standardized open-ended interview, closed quantitative interview. Their characteristics, strengths and weaknesses are outlined in Figure 7.8.

An interview has the advantage of being flexible, of eliciting spontaneous responses, and providing the opportunity for both questions and responses to be clarified if necessary. It is particularly suited to exploring issues in-depth, and, in some instances, to dealing with sensitive matters. This is in contrast to self-administered questionnaires in which respondents can experience difficulties in interpreting questions, writing down their responses or in fitting their answers into pre-set response categories (Keats, D. 1988: 12-13).

The main disadvantages of interviewing is the potential for the presence of the interviewer to influence the responses of the interviewee, and for interviewer variability in administering the questions. The principles discussed above in relation to written questionnaires also apply to face-to-face and telephone interviews.

b. Telephone Interviews

Telephone interviews are less costly and time-consuming than face-to-face interviews, and may be less prone to interviewer bias. However, they may not be as effective for dealing with certain kinds of sensitive matters. A particular difficulty associated with telephone interviews is that they rely on verbal communication alone and respondents must retain the information given in the questions and response categories. To deal with the problem of limited retention, response categories should be reduced and questions simplified, and/or divided into two parts where possible (de Vaus, 1990:92).

Further Reading:

Education Department of Victoria, Curriculum Branch (1985) Destination Decisions Decision-making strategies for School Communities, Education Department of Victoria, Ch.11.

Keats, D. (1988) Skilled Interviewing. The Australian Council for Educational Research Ltd., Hawthorn, Victoria.

Lavrakas, P.J. (1987) Telephone Survey Methods Sampling, Selection and Supervision, Sage Publications, Newbury Park, California.

Patton, M.Q. (1987) How to Use Qualitative Methods in Evaluation, Sage Publications, Newbury Park, California. Ch. 5.

c. Group Interactions

Group interactions allow for the exploration of a range of views on a particular issue or program. They enable the evaluator to tap the perceptions of a number of people at the one time and therefore, for a given sample size, are less costly than individual interviews. In contrast to the one-to-one interview, there may be less bias due to interviewer presence, although undue influence may be exerted by more dominant members of the group. Three types of group interactions are discussed here: focus groups, nominal groups and search conferences.

(i) Focus Groups

The focus group interview generally involves eight to 12 individuals who discuss a particular topic in response to questions posed by a moderator. The term 'focus' implies that the interview is limited to a small number of issues, and typically lasts from one and half to two hours (Stewart, & Shamdasani, 1990:10). The purpose of the interview is to gather information in a social context where participants can consider their own views in the context of the views of others. It is neither a decision-making nor a problem-solving forum, nor is there any attempt to reach consensus. (Patton, 1978:135)

Focus groups can be useful at any stage in an evaluation. For example:

- as part of a needs assessment with both experts in the field and potential clients
- during a program to identify key elements in a program's implementation, and program strengths and weaknesses
- as a follow-up to the program to gather perceptions of outcomes (Patton, 1978:136).

Focus group interviews offer the following advantages when used for program evaluation purposes:

- information can be collected from a number of respondents at the one time which means that the overall number involved in the sample can be increased significantly
- there is a tendency for the most important issues to emerge through group discussion and shared views

are relatively easy to identify

- they are usually enjoyed by the participants.

There are, however, a number of disadvantages:

- the number of questions that can be posed is limited as time must be allowed for all participants to respond and for discussion to occur. For example, Patton suggests that it is not possible to ask more than ten questions in a group of eight people in an hour.
- good moderating skills are required to avoid the discussion being dominated by one or two individuals and to draw out the views of quieter participants (Patton, 1978: 136-137).

In contrast to the traditional moderator-led focus group interview, is the synergetic focus group discussion which allows participants to select issues that are most pertinent for them. 'The synergetic focus group discussion is initiated by a monologue given by the researcher who provides a focus for the discussion by presenting the topic to be discussed and identifying a variety of potential perspectives on the topic. The purpose of the monologue is to create a permissive environment where participants feel comfortable discussing any aspect of the topic, confident that the researcher is open to their opinions. Once the monologue has been given the researcher takes no further part in the discussion .... After the participants have completed their discussion, the researcher may ask participants to clarify any aspects of the discussion which were unclear and may also raise any other issues which did not emerge in the discussion.' (Russell & Lidstone, 1993)

This technique is particularly effective in collecting 'inside' information from clients about their perceptions of a program or service.

Further Reading:

Krueger, R.A. (1988) Focus Groups: A Practical Guide for Applied Research, Sage Publications, California.

Stewart, D.W. & Shamdasani, P.N., (1990) Focus Groups Theory and Practice, Sage Publications Inc. Newbury Park, California.

Russell, A. & Lidstone, J. (1993) 'The Library through student Eyes; Evaluation using Synergetic Focus Group Discussions', Conference Paper, Australasian Evaluation Society International Conference, June, 1993.

(ii) Nominal Groups

Nominal Groups are used when problem-solving, setting priorities and/or decision-making is required. It is achieved with minimal discussion and involves the following major steps:

- generation of ideas. A group of five to 10 people is required to make individual written responses to a question posed at the beginning of the session. For example, 'What problems do we have with our workplace basic education program ?'

- presentation of ideas. Each person gives one idea which is written down on a large sheet of paper by a group recorder.

- clarification of ideas. Participants are encouraged to give reasons for and against the ideas, but argument is discouraged.

- voting on the ideas. Participants are asked to select items from the list and give them a priority rating. The votes are then collected and tallied.

- discussion of the resulting priorities. If needed, a final vote possibly using more refined voting techniques (McKillop, 1987).

The nominal group technique is a structured process designed to overcome some of the common problems associated with group dynamics i.e. the tendency for some individuals to dominate, for idea evaluation to discourage idea generation, and moving away from the agenda (McKillop, 1987: 88). However, it requires a skilled facilitator, is relatively time-consuming in that it may take at least one and a half hours to complete the process, and is not appropriate for simple problems.

Further Reading:

Education Department of Victoria, Curriculum Branch (1985) Destination Decisions, Decision-making Strategies for School Communities, Education Department of Victoria, Ch.12.

McKillop, J.(1987) Need Analysis Tools for the Human Services and Education, Sage Publications, Newbury Park, California, Ch. 6.

(iii) Search Conferences

'A search conference is an event designed to give direction to a projected program or policy. It involves program staff and administration in a search for appropriate solutions given their knowledge of the context in which the program is to be developed and delivered.... It is based on the assumption that people want to create their own futures, that they will act creatively and purposefully to achieve this end.'(Owen, 1993:162)

Search conferences require meticulous planning and need to be well managed. They address questions such as 'Where have we come from, what have we done well, what have we done badly, where do we want to go and what are the factors important in reaching our 'desired future' ? (Emery, 1990: 261)

Further Reading:

Emery, M. (1982) Searching, Canberra. Centre for Continuing Education, Australian National University.

Emery, M. (1990) 'The Search Conference as Evaluation Planning', Proceedings of the National Evaluation Conference, Vol. 1, Australasian Evaluation Society, Sydney, pp.259-262.

3. Information Compiled by Using Media

a. Audio-Tape

Tape-recorders can be used in a variety of settings but are particularly valuable when conducting open-ended interviews and where it is important to capture the actual words spoken by the interviewee. A primary justification for audio-taping is that it allows the interview to flow freely in conversational style, permitting the interviewer to attend more closely to the interviewee, to focus on opportunities to probe for further responses and to think ahead about new questions that might be relevant.

The use of the tape-recorder should generally not be used as a total substitute for taking notes. Brief notes in the form of major points and key phrases help the interviewer formulate questions, refer back to statements made previously and can later be used to locate relevant sections on the audio-tape (Patton, 1987:137-138).

It should be noted that the technical quality of the tape-recording is critical and that full transcription of interviews is extremely time-consuming and costly.

b. Video-Tape

Video-tapes may be used in a wide variety of settings and provide a permanent record of 'segments of the action'. For example, a video-taping of different classes may reveal similarities or differences over time or across subjects (Education Department of Victoria, 1985: 81).

Video-taping is associated with a number of disadvantages. Video-tapes are:

- intrusive and may affect participant behaviour
- selective, in that they have a limited field of vision and are often edited later
- costly, time-consuming and require special skills.

c. Time-Lapse and Still Photography

Time-lapse and still photography offer similar advantages to video-taping, although are even more selective in the 'segments of action' which they capture. They are less intrusive and less costly than video-taping, but their use may still affect participant behaviour depending on the context. Nevertheless, there have been some impressive uses of photography in evaluation (e.g. Hurworth & Sweeney, 1992). One advantage is that, in addition to using photographs as a form of data collection, they can also be used to improve the presentation of an evaluation report.

4. Unobtrusive Techniques

A major difficulty in collecting data in the field is the potential for the data collection techniques or instruments themselves to influence the reaction of participants and affect the normal operations of the program. Where participants know that they are being observed, they may behave and respond differently than under normal circumstances and hence findings may be distorted.

If there is a need for the evaluation to interfere minimally with the program under review, the use of unobtrusive measures may be necessary. Some useful measures of this kind include:

- physical traces (e.g. writing left on a white-board, the amount of wear on a carpet as a measure of the number of people interested in a particular museum exhibit)

- archival records (e.g. attendance records, publications, referrals, program change requests, etc.)
- private records (e.g. diaries, essay drafts, etc.)
- simple observation i.e. observation in settings where participants are not aware that they are being observed
- contrived observation (e.g. unobtrusive video-taping) (Education Department of Victoria, 1985: 84-85).

Further Reading:

Webb, E.J., Campbell, D.T., Schwartz, R.D. & Sechrest, L. (1966) Unobtrusive Measures: Non-Reactive Research in the Social Sciences, Rand McNally, Chicago.

5. Existing Records

Using existing records for data collection has the advantage of being unobtrusive (as indicated in the previous section) but may be difficult in that they are often incomplete, disorganised and hard to decipher (particularly if hand-written).

Examples of existing records include:
- certificates upon completion of activities
- completed student workbooks
- products produced by students
- work portfolios
- logs, journals and diaries
- progress charts and checklists
- report cards
- posters
- standardized test scores
- attendance and enrolment logs
- research reports
- annual reports
- circulars
- legal documents
- mileage records
- timetables
- letters of support or complaint
- budgets and costings.

References

Australian Committee for Training Curriculum (1993) National Framework of Adult English Language, Literacy and Numeracy Competence, ACTRAC Products Ltd., Frankston, Victoria.

de Vaus, D.A. (1990), Surveys in Social Research (2nd ed.) Allen & Unwin Australia Ltd., North Sydney, NSW.

Dillman, D.A. (1978) Mail and Telephone Surveys: the total design method, New York, Wiley.

Dixon, J. & Harding, G. (1990) 'Delphi Forecasting Methodology', Workshop Paper, Conference Proceedings, National Evaluation Conference, AES.

Education Department of Victoria, Curriculum Branch, (1985) Destination Decisions Decision-making Strategies for School Communities, Government Printer, Melbourne.

Emery, M. (1990), 'The Search Conference as Evaluation Planning', Proceedings of the National Evaluation Conference, Vol. 1, Australasian Evaluation Society, Sydney, pp.259-262.

Fowler, F (1984) Survey Research Methods, Sage Publications, Beverly Hills, CA.

Gallup, G.H. (1947) 'The Quintamensional Plan of Question Design', Public Opinion Quarterly 11, pp. 385-393.

Hurworth, R. & Sweeney, M.(1992) 'Australasian Evaluation Society 1992 International Conference: The Use of Photography in Evaluation', Evaluation Journal of Australasia, Vol.4, No.2.pp.29-42.

Keats, D (1988) Skilled Interviewing, ACER, Hawthorn, Vic.

King, J.A., Morris, L.L. & Fitzgibbon, C.T. (1987) How to Assess Program Implementation, Sage Publications, Newbury Park, California.

McKillop, J. (1987) Need Analysis Tools for the Human Services and Education, Sage Publications, Newbury Park, California.

Owen, J.M. (1993) The Potential of Talking Books for Rural Schools in Victoria. A Report to the Country Education Project, Centre for Program Evaluation, The University of Melbourne.

Patton, M. Q (1978) Utilization-Focused Evaluation (1st edition), Sage Publications, Beverly Hills, CA.

Patton, M. Q (1987) How to Use Qualitative Methods in Evaluation, Sage Publications, Newbury Park, CA.

Russell, A. & Lidstone, J. (1993) 'The Library through Student Eyes; Evaluation using Synergetic Focus Group Discussions', Conference Paper, Australasian Evaluation Society International Conference, June, 1993.

Stanley, L.A. (1987) Training Need Assessment, International Center for Public Enterprises in Developing Countries, Ljubljana, Yugoslavia.

Stewart, D.W. & Shamdasani, P.N., (1990) Focus Groups Theory and Practice, Sage Publications Inc. Newbury Park, California.

FIGURE 7.1: RATING SCALES: EXAMPLES

A. Verbal Likert-Style Format

Statement: The existing child minding facilities at our community house are adequate for my current needs

☐	☐	☐	☐	☐
Strongly Agree	Agree	Can't Decide	Disagree	Strongly Disagree

B. Diagramatic Likert-Style Format

Statement: The existing child minding facilities at our community house are adequate for my current needs

Strongly Agree 1 2 3 4 5 6 7 8 9 10 Strongly Disagree

FIGURE 7.2: GUIDELINES FOR CONDUCTING A DELPHI STUDY

HOW DO YOU USE IT?

1. An individual or a group chooses an issue to be decided through the Delphi. The issue is phrased as a question, e.g.: 'What should be done to improve the quality of education at this school?' Check that the question is easily understood.
2. Decide how many people will be involved. This will vary according to the purpose of the Delphi. A Delphi designed to provide expert judgements about how to reach certain goals may involve as few as say, fifteen educational experts. On the other hand, a Delphi intended to develop school goals should involve at least fifty people so that a wide range of views is represented. In some cases it is best to involve everyone who wants to take part. This will also ensure that there will be enough people left to continue the Delphi if some drop out. Some Delphi have used more than 400 respondents. Remember though, the more respondents, the more work for the Delphi managers. It is worth nothing that past experiences shows that a Delphi using 400 people doesn't produce many more ideas than a Delphi involving fifty people.

   When a large-scale Delphi is impractical, but organisers want to involve as many people as possible, the solution might be to run a Delphi involving about fifty people and then use their answers to develop a questionnaire which could be answered by large numbers of people.
3. Contact the people and invite them to take part in the Delphi. Explain how the Delphi works, tell them why you are running it and how you intend to use the results.

   Assure them that there will be no meetings to attend and that anonymity is assured by the process. It is possible to ignore this step and proceed directly to the next.
4. Write a covering letter to be mailed out with the Round 1 question.

   The letter should:
   - thank people for agreeing to take part in the Delphi;
   - explain why help is needed;
   - describe the Delphi procedure and how the results will be used;
   - give a date for the return of responses (two weeks from the mailing date).

   Include a stamped, addressed envelope for the Round 1 replies. If the covering letter is your first contact with the person, it should also include the points listed in step 3. The letter should be typed and should not be more than one page long.
5. Send out the Round 1 questionnaire with the covering letter. Invite people to list any ideas they have that are relevant to the question. Each idea should be written on a new line. This saves work later on.
6. Send out reminder notices. You will find that some people need extra encouragement. Two forms of encouragement are common:
   - A letter posted one week after the initial mailing to remind everyone of the response date and to thank those who have already sent in their answers.
   - A phone call to each person to ask if there are any problems. Use the opportunity to remind people of the response date and the importance of their views.
7. Collect all the Round 1 responses and write each item on a separate small card. Alternatively, simply cut up the response lists so that you separate the items. Sort the items, making different heaps of items that are similar. Number the heaps. Read the items in heap one and write a straightforward sentence that sums up the view conveyed by the various separate contributions. Work through each heap of cards until you have a set of sentences which sum up the answers to Round 1.
8. Prepare the task for Round 2. Usually respondents are given copies of the sentences written in step 7. They are invited to add new items or to briefly comment on the items listed. They are then asked to vote for the items they believe to be most important. Delphi managers will need to decide the voting method to be used, e.g.:
   - a rating, e.g. a Likert scale (see 'Questionnaires');
   - a ranking, e.g. the respondent could be asked to rank his or her top five selections; or
   - Q-sort (see 'Questionnaires').
9. Write a covering letter for Round 2. It should:
   - thank those who responded in Round 1;
   - point out the importance of their continued involvement;
   - explain Round 2;
   - give a date for return of responses (two weeks from the mailing date). A stamped, addressed envelope should again be included for Round 2 answers.
10. Send out Round 2 questionnaire and covering letter.
11. A week later send a reminder letter or ring each person.
12. Collect the Round 2 responses.

    Tally the votes for each item.

    Summarise the comments made about each item.

    Add any new items to the list of summaries.
13. Develop a questionnaire for Round 3 by listing the Round 2 items plus any new items. Show how many votes each item received in Round 2. Summarise the comments made in Round 2.

    Round 3 may be a ranking task, a rating task or a Q-sort task.
14. Write a covering letter for Round 3.
15. Send out Round 3 materials.
16. Send out reminder notices or make the necessary phone calls.
17. Collect the Round 3 responses and prepare a report. If necessary, the Delphi process may be extended to a fourth or fifth round. In most cases, there is a clear consensus after Round 3 (or Round 4, at worst).

    The final report should be a summary of each round and the overall results of the Delphi. As everyone is sent a copy of the report, this is a good opportunity to thank respondents for their help.

Source: Curriculum Branch, Education Department of Victoria (1986) Destination: Decisions, Decision-making Strategies for School Communities, Education Department of Victoria, p.48-49.

BEST COPY AVAILABLE

FIGURE 7.3: RANKING SCALE : AN EXAMPLE

How important are the following characteristics for a good literacy program? (Please list the items in order of importance e.g. if 'learners participating in setting objectives' is most important, place 1 in the box next to it. For the second most important, place 2 in the box and so on.)

☐ Reading, writing and numeracy skills are presented as parts of an integrated process

☐ Learners participate in setting objectives for sessions.

☐ Learners' prior knowledge and cultural experience is built upon in sessions.

☐ Accurate and regular records are kept of learner progress.

☐ Session planning is guided primarily by learner needs and interests.

FIGURE 7.4: SEMANTIC DIFFERENTIAL FORMAT: AN EXAMPLE

How would you describe your trainers in this program?

Friendly	1	2	3	4	5	6	7	Unfriendly
Competent	1	2	3	4	5	6	7	Incompetent
Outgoing	1	2	3	4	5	6	7	Outgoing
Fair	1	2	3	4	5	6	7	Unfair

FIGURE 7.5: DIFFERENT RESPONSE FORMATS: AN OVERVIEW

Figure 7.5a: An Example of An Open-Response Questionnaire

DOCUMENTATION QUESTIONNAIRE
Peer-Tutoring Program

The following are questions about the peer-tutoring program implemented this year. We are interested in knowing your opinions about what the program looked like in operation. Please respond to each question, and feel free to write additional information

(1) How was the peer-tutoring program structured in your classroom?

(2) How were tutors selected?

(3) How were students selected for tutoring?

(4) What materials seemed to work best in the peer-tutoring sessions? Why?

(5) What were the strengths of the peer-tutoring program this year?

(6) What changes would you make to improve the program next year?

Figure 7.5b: An Example of a Closed-Response Questionnaire

DOCUMENTATION QUESTIONNAIRE
Peer-Tutoring Program

Please answer the following questions by placing the letter of the most accurate response on the line to the left of the question. We are interested in finding out what the project looked like in operation during the past week, regardless of how it was planned to look. If more than one answer is true, answer with as many letters as you need.

____ 1. On the average, how many times did tutoring sessions take place in your classroom?
- a) never
- b) 1 or 2 times
- c) 3 or 4 times
- d) 5 or more times

____ 2. What was the average length of a tutoring session?
- a) 5-15 minutes
- b) 15-25 minutes
- c) 25-45 minutes
- d) longer than 45 minutes

____ 3. Where in the school did tutoring usually take place?
- a) classroom
- b) sometimes classroom, sometimes other room
- c) library
- d) room other than classroom or library

____ 4. Who were the tutors?
- a) only fast students
- b) fast students and some average students
- c) only average students
- d) other

____ 5. On what basis were tutees selected?
- a) reading achievement
- b) teacher recommendations
- c) general grade average
- d) other

____ 6. What materials were used by teachers and tutors?
- a) whatever tutors chose
- b) specially constructed games
- c) 'Read and Say' workbooks
- d) other

____ 7. How typical of the program as a whole was last week, as you have described it here?
- a) just the same
- b) almost the same
- c) some aspects not typical
- d) not typical at all

Figure 7.5c: An Example of a Questionnaire including both Open- and Closed- Response Formats

DOCUMENTATION QUESTIONNAIRE
Peer-Tutoring Program

The following are statements about the peer-tutoring program implemented this year. We are interested in knowing whether they represent an accurate statement of what the program looked like in operation. For this reason, we ask that you indicate, using the 1 to 5 scale after each statement, whether it was "generally true", etc. Please circle your answer. If you answer seldom or never true, please use the lines under the statement to correct its inaccuracy.

	always true	generally true	seldom true	never true	don't know
1. Students were tutored three times a week for periods of 45 minutes each. ______ ______	1	2	3	4	5
2. Tutoring took place in the classroom, tutors working with their own classmates. ______ ______	1	2	3	4	5
3. Tutors were the fast readers ______ ______	1	2	3	4	5
4. Students were selected for tutoring on the basis of reading grades. ______ ______	1	2	3	4	5
5. Tutoring used the "Read and Say" workbooks. ______ ______	1	2	3	4	5
6. There were no discipline problems. ______ ______	1	2	3	4	5

Source: King, J.A., Morris, L.L. & Fitzgibbon, C.T. (1987) How to Assess Program Implementation, Sage Publications, Newbury Park, California, p.78-80.

BEST COPY AVAILABLE

FIGURE 7.6: GUIDELINES FOR QUESTIONNAIRE LAYOUT

Answering procedures

With open-ended questions ensure that you leave sufficient space for answer to avoid people cramming responses but do not leave so much as to discourage completing the questionnaire because of the time it will apparently take.

With closed questions people can be asked to either tick appropriate boxes or brackets or circle a number next to responses (see Figure 6.2)

When using any of these procedures, the area for answering can be on the left or right of the response but make sure you justify your typing on the answer side as below.

1 [ ] Agree		Agree [ ] 1
2 [ ] Disagree	OR	Disagree [ ] 2
3 [ ] Can't decide		Can't decide [ ] 3

When a series of Likert-style questions are to be used it can be efficient to present them in a matrix format (see Figure 6.3).

Contingency questions

Since you do not want respondents to waste time reading questions which are not relevant to them we can use filter or contingency questions as in Figure 6.4.

The use of arrows and inset boxes to highlight follow-up questions is a useful way of avoiding confusion when using contingency questions.

Instructions

To provide flow, use the following types of instruction where appropriate.

1. *General Instructions:* These should include an introduction to the purpose of the questionnaire, assurance of confidentiality, how the respondent was chosen, how and when to return the questionnaire (where relevant).
2. *Section introductions:* When the questionnaire can be divided into subsections provide a brief introduction to each section such as 'Finally we would like to know just a little about your background so we can see how different people feel about the topics about which you've answered questions.'
3. *Question instructions:* Indicate how many responses the respondent can tick (e.g. the most appropriate, as many as apply, one only).
4. *'Go to' instructions:* Ensure you make use of these when using contingency questions.

Figure 6.2 Different answering formats for closed-choice questions

Square brackets, parentheses or boxes (tick the box)

1 [ ] Agree		1 ( ) Agree		1 Agree
2 [ ] Disagree	OR	2 ( ) Disagree	OR	2 Disagree
3 [ ] Can't decide		3 ( ) Can't decide		3 Can't decide

Precoding (circle the number)

1 Agree
2 Disagree
3 Can't decide

Figure 6.3 A matrix presentation of Likert-style questions

	Strongly agree	Agree	Disagree	Strongly disagree	Don't know
This country has a rich cultural life in music, art, and literature	[ ] 1	[ ] 2	[ ] 3	[ ] 4	[ ] 5
Unfortunately this country is a long way from the centre of things.	[ ]	[ ]	[ ]	[ ]	[ ]
Only a small proportion of young people can get a decent education in this country.	[ ]	[ ]	[ ]	[ ]	[ ]
This country is a good place for a person to start a small business.	[ ]	[ ]	[ ]	[ ]	[ ]

Figure 6.4 An illustration of contingency questions

Were you born in Australia?

1. [ ] Yes (go to question 2)
   [ ] No

(a) In what country were you born? ______________
____

(b) For how many years have you lived in Australia?
_____ ____ yrs.

(c) Are you an Australian citizen?
[ ] Yes
[ ] No

Now go to Question 2

Use of space

To encourage people to complete a questionnaire avoid cluttering it. The following hints may help:

1. Print questions on one side of the page only. It is too easy for people to miss questions printed on the backs of pages. The blank backs of pages also are useful for respondents to write additional comments.
2. Provide a column about 2.5 centimetres wide on the right hand side for computer coding (see Chapter 14).
3. Leave sufficient space for open-ended questions.
4. List alternative responses down rather than across the page.

Order of questions

A good questionnaire is one in which there is a good logical flow to questions. The following points provide some guidelines.

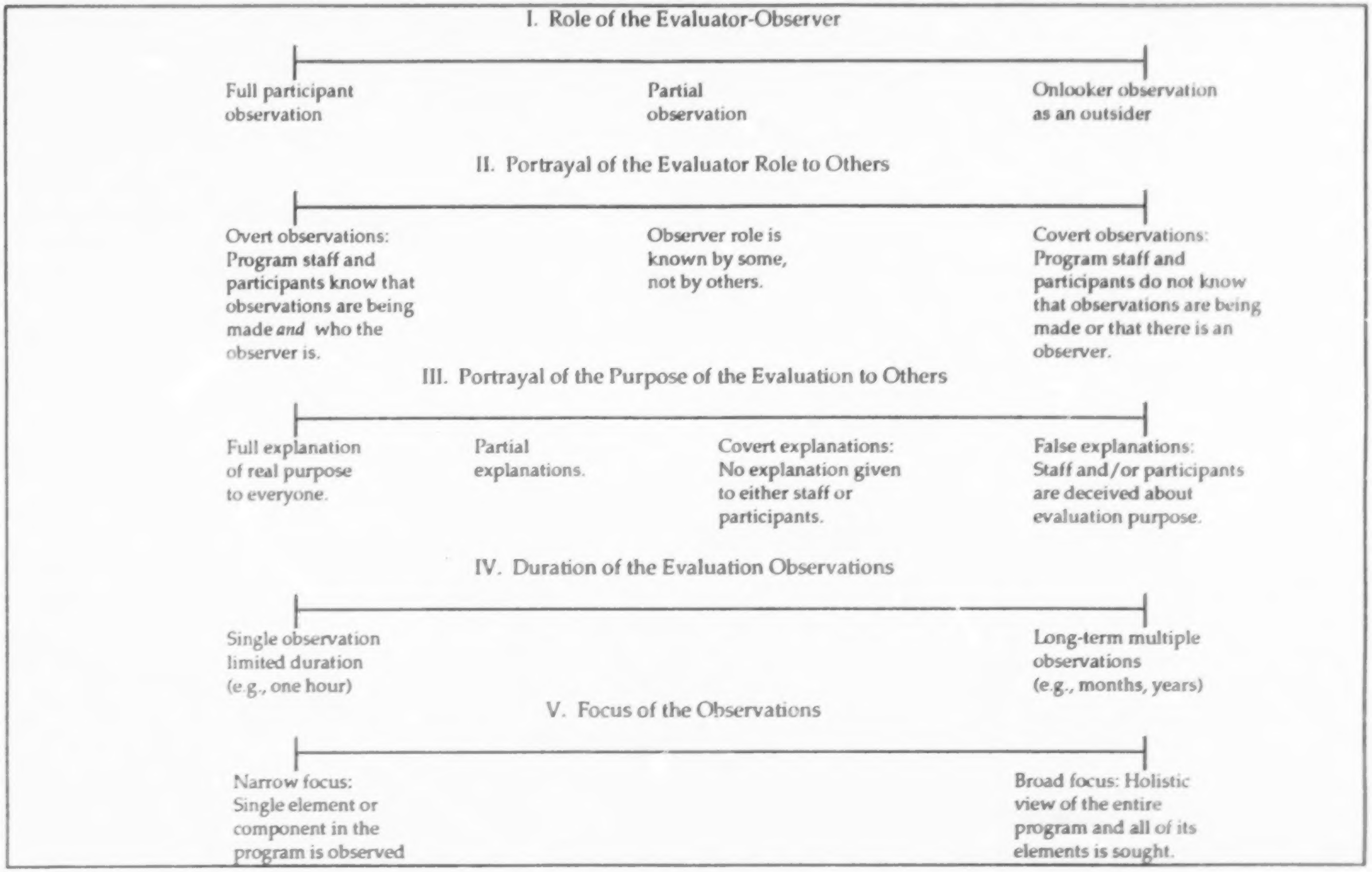
1. Commence with questions the respondent will enjoy answering.
   a. There should be easily answered questions.
   b. Factual questions should be used initially.
   c. Do not start with demographic questions such as age, marital status, etc.
   d. Ensure that the initial questions are obviously relevant to the stated purpose of the survey.
2. Go from easy to more difficult questions.
3. Go from concrete to abstract questions.
4. Open-ended questions should be kept to a minimum and where possible placed towards the end of the questionnaire.
5. Group questions into sections. This helps structure the questionnaire and provides a flow.
6. Make use of filter questions to ensure that questions are relevant to respondents.
7. When using a series of positive and negative items to form a scale, mix up the positive and negative items to help avoid an acquiescent response set.
8. Where possible try to introduce a variety of question formats so that the questionnaire remains interesting.

Setting up for coding

If the data are to be analysed by computer it is useful to prepare for this by allocating codes to responses in the questionnaire so that a number is printed in the questionnaire next to responses.

Source: de Vaus, D.A. (1990) Survey in Social Research, 2nd Ed. Allen & Unwin Australia Ltd., North Sydney, NSW, p.89-93.

FIGURE 7.7: FIVE DIMENSIONS ALONG WHICH FIELDWORK VARIES


I. Role of the Evaluator-Observer
Full participant observation
Partial observation
Onlooker observation as an outsider
II. Portrayal of the Evaluator Role to Others
Overt observations: Program staff and participants know that observations are being made *and* who the observer is.
Observer role is known by some, not by others.
Covert observations: Program staff and participants do not know that observations are being made or that there is an observer.
III. Portrayal of the Purpose of the Evaluation to Others
Full explanation of real purpose to everyone.
Partial explanations.
Covert explanations: No explanation given to either staff or participants.
False explanations: Staff and/or participants are deceived about evaluation purpose.
IV. Duration of the Evaluation Observations
Single observation limited duration (e.g., one hour)
Long-term multiple observations (e.g., months, years)
V. Focus of the Observations
Narrow focus: Single element or component in the program is observed
Broad focus: Holistic view of the entire program and all of its elements is sought.

Source: Patton, M.Q. (1987) How to Use Qualitative Methods in Evaluation, Sage Publications, Newbury Park, California, p.81.

FIGURE 7.8: DIFFERENT TYPES OF EVALUATION RESEARCH INTERVIEWS: Characteristics, Strengths, Weaknesses

Type of Interview	*Characteristics*	*Strengths*	*Weaknesses*
(1) Informal conversational interview	Questions emerge from the immediate context and are asked in the natural course of things; there is no predetermination of question topics or wording.	Increases the salience and relevance of questions; interviews are built on and emerge from observations; the interview can be matched to individuals and circumstances.	Different information collected from different people with different questions. Less systematic and comprehensive if certain questions do not arise "naturally". Data organization and analysis can be quite difficult.
(2) Interview guide approach	Topics and issues to be covered are specified in advance, in outline form; interviewer decides sequence and wording of questions in the course of the interview.	The outline increases the comprehensiveness of the data and makes data collection somewhat systematic for each respondent. Logical gaps in data can be anticipated and closed. Interviews remain fairly conversational and situational.	Important and salient topics may be inadvertently omitted. Interviewer flexibility in sequencing and wording questions can result in substantially different responses, thus reducing the comparability of responses.
(3) Standardised open-ended interview	The exact wording and sequence of questions are determined in advance. All interviewees are asked the same basic questions in the same order.	Respondents answer the same questions, thus increasing comparability of responses; data are complete for each person on the topics addressed in the interview. Reduces interviewer effects and bias when several interviewers are used. Permits decision-makers to see and review the instrumentation used in the evaluation. Facilitates organisation and analysis of the data.	Little flexibility in relating the interview to particular individuals and circumstances; standardised wording of questions may constrain and limit naturalness and relevance of questions and answers.
(4) Closed quantitative interview	Questions and response categories are determined in advance. Responses are fixed; respondent chooses from among these fixed responses.	Data analysis is simple; responses can be directly compared and easily aggregated; many questions can be asked in a short time.	Respondents must fit their experiences and feelings into the researcher's categories; may be perceived as impersonal, irrelevant, and mechanistic. Can distort what respondents really mean or experienced by so completely limiting their response choices.

Source: Patton, M.Q. (1987) How to Use Qualitative Methods in Evaluation, Sage Publications, Newbury Park, California, p.116.

Chapter 8 COLLECTING AND ANALYSING EVALUATIVE INFORMATION

The topics covered in this chapter bring the reader into the realm of sampling techniques, probability theory and statistical analysis. The aim here is to introduce key concepts within these areas showing how they relate to evaluation questions, information collection and analysis. Those seeking a detailed understanding of the concepts will need to refer to the further reading lists provided.

A. Selecting and Reviewing Information Collection Alternatives

As indicated in the *overview framework* in Figure 2.3, the selection of information collection techniques depends on :

- the key evaluation issues
- when the evaluation findings are required
- the availability of resources and expertise.

For each evaluation issue, a combination of different data collection methods should be used to strengthen one's confidence in the validity of the findings.

B. Reviewing Information Collection Techniques

Figure 2.3 provides a series of useful questions when reviewing the appropriateness of information collection techniques. These are:

- will the data collected give a comprehensive picture of what is being evaluated?
- does the data management plan make effective use of existing data?
- will the cost of data collection be justified, given the amount and kind of information it will provide?
- will the information be reliable?
- can the data collection be carried out without unduly disrupting the program and taking too much of the time of the program providers?
- are the data collection procedures legal and ethical?
- can the data be collected and analysed within the time constraints of the study?
- is a pilot study necessary to address some of the questions identified above?

C. Testing Information Collection Techniques Via A Pilot Study

A pilot is a small scale study designed to test the effectiveness and efficiency of selected data collection techniques and instruments. This involves trying out the instruments under conditions similar to those expected in the evaluation, and so it should include a range of people similar to those who will participate in the evaluation. The pilot could, for example, be used to test the design of interview schedules, self-report instruments or surveys intended for widespread distribution. It is carried out prior to the evaluation proper to ensure that resources are not wasted due to unanticipated difficulties in implementing the data collection strategy and/or the collection of unnecessary or invalid data. An example of this use is demonstrated in the following DEET Case Study (Challender, 1993).

> DEET Case Study: The Pilot of a National Survey of Employer Satisfaction with the Commonwealth Employment Service (CES).
>
> 'Recently (April/May 1993), DEET designed a pilot study to test the most effective survey techniques for determining the extent and nature of employer satisfaction with the Commonwealth Employment Service (CES).'
>
> The pilot survey involved a sample population of around 2,500

employers. Stratification of the sample population was by state, industry type, establishment size, and where possible, by geographic split (metropolitan and non-metropolitan). Three main groups were targeted:

- active CES users i.e. employers who were active CES users in the 12 months prior to the survey
- employers who had been approached by CES in the past year and who were therefore aware of their services, but had not recently used them
- employers who did not use CES and had not been approached by them about their services in the past year.

'For the first two groups DEET provided the successful consultant with listings of employers selected from the DEET JOBSYSTEM database. JOBSYSTEM is a database designed to assist Regional Office staff to deliver and monitor Government policies and programs. JOBSYSTEM has information on the number and characteristics of clients, statistical reports, and details about project activities.

For the third group however, as a condition of the project brief, the consultant was asked to identify 6,000 additional employers from the computerised version of the Telecom Yellow Pages in order to obtain the 500 sample listing of non-CES users needed to complete the third survey population. Listings supplied by the consultants were made available in a predetermined format to enable matching against the DEET database (JOBSYSTEM) to ensure selected employers did not fall into the two categories.'

One of the main objectives of the pilot National Employer Survey was to test the adequacy of the :

- proposed sampling frame (i.e. the population listing used for obtaining the sample)
- survey methodology (telephone and mail methods)
- questionnaire design.

a. Limitations of the Sampling Frame

Problems associated with the population frames of both JOBSYSTEM and the computerised version of the Telecom Yellow Pages were identified. For example:

Using JOBSYSTEM, it was discovered that there were duplicate records and some employers were recorded more than once. This meant that information regarding 'date of last contact' and 'date of last vacancy' could be incorrect for some employers selected and result in incorrect classification into sample populations.

Using Telecom Yellow Pages, it was found that descriptive information about the size of operations was not available. This meant that a telephone screening process using a classificatory question as a quota system was necessary in order to include the correct number of respondents in each employer size category.

This made it clear that an alternative source of listings of employers would be more appropriate for the larger survey.

b. Survey Methodology

Both telephone and mail surveys were trialled in order to determine the most effective way of contacting the right person within companies/ businesses to respond to specific recruitment issues.

Those in the sample population were approached either by:

- telephone interview - using Computer Assisted Telephone Interviewing (CATI). The CATI system involved programming a survey questionnaire directly into a personal computer. The telephone interviewer then read the questionnaire from the monitor and recorded the answers directly into the computer; or
- mail questionnaire, with a single CATI follow-up.

The findings from the pilot study showed that telephone interviews had a number of advantages over mailed questionnaires:
- screening methods ensured that the right person was contacted
- the purposes of the project could be explained directly and this encouraged cooperation
- interim reports from the CATI system were available on a weekly basis because data analysis was virtually instantaneous
- the initial response rate for the mailed questionnaires yielded a low 24% response rate, which was doubled or trebled in most States as a result of a telephone follow-up phase.

Overall, it was concluded that the CATI methodology was more effective.

c. Questionnaire Design

'The pilot survey was preceded by a small pre-test of both the telephone and mail questionnaires. The pretest was conducted using the consultant's CATI system. In total 90 pretests were conducted. These included 10 CATI interviews and 10 questionnaires from each of the three employer sizes and each of the targeted groups.

Following the pretest a debriefing with the consultants was held. The main objective of this meeting was to review outcomes from the pretest and to revise the CATI and self completion questionnaire for the pilot survey. Pretesting highlighted areas of confusion. Terminology prone to misinterpretation was identified and eliminated. It also allowed for other issues to be finalised before the pilot such as questionnaire layout, question code frames and procedures.

The pretests identified potential problems in wording or questionnaire mechanics but could provide little information related to the content of the questions and answers. However, with the larger pilot study, the distribution of answers could be tabulated. Distributions that were markedly different from what was desired or expected suggested areas for revision. In addition, patterns of association could be examined. Analysis of correlations gave some indication of whether or not questions were measuring what was expected.' (Fowler, 1984:104-105)

The pilot study of the National Employer Survey provided the opportunity to reduce significant sources of bias prior to the larger scale survey being conducted.

Further Reading:

Fowler, F. (1984) Survey Research Methods, Sage Publications, Beverly Hills, California.

Suchman, E. (1969) Evaluative Research, Russell Sage Foundations, USA.

D. Finding A Sample

'One way of finding out about a group of people is to collect information from everyone in the group. For large groups of people this is prohibitively expensive and impractical. The alternative is to collect information from only some people in the group in such a way that their responses and characteristics reflect those of the group from which they are drawn. This procedure is much cheaper, faster and easier than surveying all members of a group. This is the principle of sampling.'

'A sample is obtained by collecting information about only some members of the population. Samples can reflect the populations from which they are drawn with varying degrees of accuracy. A sample which accurately reflects its population is called a representative sample.'(de Vaus, 1990:60)

'Sampling can provide an efficient and accurate way of obtaining information about a large number of cases. Just how efficient and accurate depends on the type of sample used, the size of the sample and the method of collecting data from the sample. In the end decisions about samples will be a compromise between cost, accuracy, the nature of the research problem and the art of the possible.' (de Vaus, 1990:79)

Two basic types of sampling are probability and non-probability sampling:

a. Probability Samples

A probability sample is one in which each person in the population has an equal, or at least a known, chance (probability) of being selected. Equal probability of selection is gained using the principle of random selection. This involves establishing a sampling frame which is a listing of all members of the population and then, in effect, pulling their names out of a hat.

There will be differences between the sample and the sampling population that will occur by chance alone. However, it is possible to use probability theory to estimate how closely the sample population is likely to reflect the total population.

While probability samples are often preferable to non-probability samples (see below) in that they are likely to be more representative and estimates of accuracy can be done, they are not always practical or necessary.

Four common forms of probability sampling include:

- Simple Random Sampling

  In simple random sampling, members of a population 'are selected, one at a time, independently of one another, without replacement; once a unit is selected, it has no further chance to be selected.' (Fowler, 1984:23)

  For example: 'If there were 8,500 people on the list, and the goal was to select a simple random sample of 100, the procedure would be straightforward. People on the list would be numbered from 1 to 8,500. Then a computer, a table of random numbers, or some other generator of random numbers would be used to produce 100 different numbers in the range from 0001 to 8,500. The individuals corresponding to the 100 numbers chosen would constitute a simple random sample of that population of 8,500.' (Fowler, 1984:23)

- Systematic Sampling

  Random sampling can often be very time-consuming unless the list of the total population is short, has all units prenumbered or can be numbered easily by computer.

  Systematic sampling can be used to simplify the process.

  'When drawing a systematic sample from a list, the researcher first determines the number of entries on the list and the number of elements from the list that are to be selected. Dividing the latter by the former will produce a fraction.

  Thus if there are 8,500 people on a list and a sample of 100 is required, 1/85 of the list is to be included in the sample; one out of every 85 persons on the list is to be selected.

  In order to select a systematic sample, a start point is designated by choosing a random number from 1 to 85. The randomized start ensures that it is a chance selection process. Given that start, the researcher proceeds to take every 85th person on the list.' (Fowler, 1984:23)

- Stratified Sampling

  At the outset of research, it is often possible to identify specific characteristics of the population to be sampled. In these instances, there is merit in 'structuring the sampling process to reduce the normal sampling variation, thereby producing a sample that is more likely to reflect the total population than a simple random sample. The process by which this is done is called stratification.'(Fowler, 1984:24) An example of stratified sampling is provided in the DEET case study described above.

- Multistage Cluster Sampling

  'When there is no adequate list of the individuals in a population and no way to get at the population directly, multistage sampling provides a useful approach.

  In the absence of a direct sampling source, a strategy is needed for linking population members to some kind of grouping that can be sampled. These groupings can be sampled at a first stage. Lists then are made of individual members of selected groups, with possibly a further selection from the created list at the second (or later) stage of sampling.' (Fowler, 1984: 26)

Selection of the sample size will depend on a number of factors:

- the degree of accuracy required
- the extent to which there is variation in the population in regard to the key characteristics of the study
- the time and resources available for the study.

Non-response can create a problem by reducing sample size. Methods to counter this include:

- drawing an initial sample larger than that required
- using well-trained people to collect the data
- employing systematic methods of following up non-respondents.

Non-response, however, also introduces bias because those who choose not to respond may well differ in significant ways from those who do. Methods to counter bias include:

- observation of characteristics of non-respondents where contact is made (e.g. sex, age, ethnic background, etc.)
- identifying characteristics of non-respondents from

sampling frames (e.g. official records)

- a comparison of the characteristics of those in the sample with the characteristics of the total population (if known) which will indicate the degree of bias.

b. Non-Probability Samples

A non-probability sample is one in which some respondents have a greater, but unknown chance than others of selection. Three types of non-probability sampling are discussed here:

- Purposive Sampling
- Quota Sampling
- Availability Sampling.

(i) Purposive Sampling

While the power of probability samples lies in their ability to permit generalisation to a larger population, the power of non-probability sampling 'lies in selecting information-rich cases for study in-depth. Information-rich cases are those from which one can learn a great deal about issues of central importance to the purpose of evaluation, thus the term 'purposeful.' (Patton, 1987:51-52) For example, if the purpose of evaluation is to improve a program's ability to cater for the literacy needs of women in particular, then it may be more illuminating to conduct an in-depth investigation of the needs, perceptions, and interests of a small group of women rather than employing a large-scale, statistically significant sample.

Figure 8.1 identifies a number of different forms of purposive sampling and describes their primary purpose (Patton, 1990:182-183).

c. Quota Sampling

'Quota sampling is another common non-probability technique aimed at producing representative samples without random selection of cases. Interviewers are required to find cases with particular characteristics: they are given quotas of particular types of people to fill. The quotas are organised so that in terms of the quota characteristics the final sample will be representative.' (de Vaus, 1990: 78) Note that quota sampling was used in the DEET pilot survey described above.

'Quota techniques are non-random because interviewers can select any cases which fit certain criteria. This can lead to bias as interviewers will tend to select those who are easiest to interview and with whom they feel most comfortable (e.g. friends). Another difficulty is that accurate population proportions may be unavailable. Finally, since random sampling is not used, it is impossible to estimate the accuracy of any particular quota sample.' (de Vaus, 1990:78)

d. Availability Sampling

This type of sampling involves taking a sample of anyone who will respond. While large sample sizes can be obtained in this way, they are in no way representative. However, they may well be useful for pilot testing questionnaires or for exploratory research designed to identify a range of views or issues.

Further Reading

de Vaus, D.A. (1990) Surveys in Social Research, Allen & Unwin Australia Pty. Ltd. North Sydney, NSW.

Fowler, F.J. (1984) Survey Research Methods, Sage Publications, Beverly Hills, London.

Moser, C. & Kalton, G. (1971) Survey Methods in Social Investigation 2nd ed. Heinemann, London.

Patton, M.Q. (1990) Qualitative Evaluation and Research Methods, Sage Publications, Newbury Park, California.

Sudman, S.(1976) Applied Sampling, Academic Press, New York.

Warwick, D.P. & Lininger, C.A. (1975) The Sample Survey: Theory and Practice, McGraw-Hill, New York.

E. Analysis of Information in Evaluation

Analysis is a fundamental aspect of any formal evaluation. It involves making sense of the collected information. It is important that the analytical processes used are defensible, meaning that they stand up to the scrutiny of stakeholders and others interested in the evaluation findings.

As indicated in the overview framework in Figure 2.3, information analysis means processing the data through the interaction of three processes:

- data reduction, the process of simplifying and transforming the raw information according to some logical set of procedures or rules.
- data display, the development of an organised assembly of information which leads to drawing of conclusions about the key evaluation questions.
- conclusion drawing, making meaning of the data in the broader context of the evaluation issue being examined.' (Owen, 1993a)

An example illustrates these points. Imagine there is interest in whether an information skills program in the workplace is having differential effects on workers from two areas of a factory. The evaluator decides that an important source of data collection in the evaluation is a competency test completed by all workers at the end of the program.

A test score for each worker is the raw data for this aspect of the evaluation. The process of **data display** involves drawing a histogram or distribution of the scores for each work area. The process of **data reduction** involves the calculation of the median and standard deviation (or spread) of the results. The data have been reduced from a set of individual scores to a set of summary statistics. In addition, a simple test (a t-test) is used to tell whether the difference between the groups is significant. On the basis of this information, the evaluator **draws a conclusion** that the program is having a differential effect on the two groups.

Analytical processes go hand in hand with data collection processes. While in the above example data analysis followed data collection, in many evaluations data analysis takes place concurrently, or in conjunction with data collection. For example, in some cases the evaluation may have a set of linked stages, the design of which depends on the findings of the previous stage. In these cases, the conclusions from one stage may set the agenda for the next stage. These could be called responsive evaluations in that the agenda for the study responds to what has been previously discovered rather than following some fixed preordinate design.

a. Evaluation as Proof: Implications for Information Analysis

Traditional evaluation techniques are aimed at proving that a given program intervention makes a difference. An elegant way to realise this is to set up a study in which one group experience the program and another parallel group do not. This style is borrowed from the agricultural paradigm of experimental and control groups and uses classical well established statistical techniques to draw conclusions about program impact.

'Comparison group designs are frequently divided into two categories: quasi- and true experimental designs (Campbell and Stanley, 1963). Both categories include comparison groups, but membership in the groups is determined in different ways. In quasi-experimental designs, individuals are assigned to experimental and control groups in such a way that differences may exist between the groups before the program begins. That means any observed differences between them after participation in the program cannot be conclusively linked to the experimental program. In true experimental designs, on the other hand, assignment of individuals results in groups that are initially as similar as possible, and any observed differences can be linked to participation in the experimental program. To guarantee the similarity of the groups for a true experimental design, you usually have to assign individuals at random to the various groups.' (Kosecoff and Fink, 1982: 85)

True experimental designs are often costly and impractical. For this reason, in naturalistic settings, quasi-experimental designs are most common, where 'control' is exercised over when the data is to be collected, the source of the data and, in some instances, the random assignment of treatments.

Further Reading:

Bhola, H.S. (1979) Evaluating Functional Literacy, Hulton Education Publications, Tehran.

de Vaus, D.A.(1990) Surveys in Social Research, Allen & Unwin Australia Pty. Ltd., North Sydney, NSW.

Pisani, R. (1985) Statistics A Tutorial Workbook, Norton, W.W. & Co., New York, London.

The use of comparison groups is suited to an impact evaluation of large-scale program interventions, for example at the regional or state level. Such an evaluation might respond to a need to establish the relative effectiveness of a set of parallel programs with a view to choosing the most effective one for future expansion to another region or state. A controlled outcomes focussed study might be part of the design. In these situations, attention must be given to the validity and reliability of the instruments used.

Instrument Validity and Reliability. An approach used to collect data is considered to be valid if it measures what it is intended to measure i.e. if the instrument used is giving you the true picture of the skill, attitude or ability it is measuring (construct validity) (Henersen, 1987:133). Examples of 'constructs' relevant to ALBE might include 'numerical ability', 'program satisfaction', or 'self-esteem'. The instrument might be a test, an attitude scale or a semantic differential.

Evidence for construct validity might be derived from several sources:

- the opinions of those qualified to make judgements. For example, curriculum specialists might be asked individually if they believe that questions related to practitioner attitudes will prompt true responses.
- correlations between the measurement of the construct and related constructs. For example, it would be expected that

the results of self-esteem measures might correlate positively with results of emotional health or achievement measures.

- criterion-group studies. These involve gathering a group of people deemed to be high or low in the construct and administering the test to them. If those people believed to have an abundance of self-esteem, for example, score very well and those believed to be low in self-esteem produce low scores, then the validity of the construct would be supported.

- an appeal to logic. It may be possible to argue that the instrument makes logical sense and will be valid so long as it is administered under the appropriate conditions (Henersen, 1987: 137-138).

An approach is considered **reliable** if the same result is obtained on repeated occasions.

Where the concept is measured by a single question, the only way to check its reliability is to ask the same people the same questions at intervals of two to four weeks and calculate the correlation co-efficient between the answers on both occasions. If the correlation is high, then it is assumed that the question is reliable. Unfortunately, because some people can recall their previous answers and it is often difficult to give the same test twice, the test re-test method has its limitations.

The best methods for testing reliability apply to measuring the reliability of scales where there is a set of questions to measure the one concept rather a single question only (de Vaus, 1990:54). Scales are frequently used in the measurement of attitudes. Attitude measurement is difficult because attitudes are not directly observable - they are abstract constructs used by humans to make sense of what people think, say and do (Henersen et.al. 1987:11). To create a scale, information contained in several specific variables is converted into one new and more abstract variable. For example, in measuring an attitude such as traditionalism, a series of questions would be used to tap into that attitude.

The development of tests and scales is a major undertaking. Where possible, an evaluator should look around for a suitable instrument rather than designing one from scratch. It is essential that, in these circumstances, evaluators check that each instrument meets validity and reliability requirements. This applies particularly to impact evaluations where the persuasiveness of the findings hinges heavily on the trustworthiness of the test or scale used to measure outcomes.

b. Evaluation as Investigation: Implications for Information Analysis

While an 'experimental' approach to program evaluation may be used in certain circumstances, there are concerns about the exclusive use of this approach in many situations. One well documented concern is that controlled designs cannot accommodate unanticipated outcomes, or indeed measure the full range of expected outcomes of a given program. At a more general level, many evaluative approaches do not focus exclusively on outcomes, or even on outcomes at all. Clearly there is a need for a more eclectic view of data collection and analysis than one limited to the preordinate design approach discussed above.

Many evaluators have moved away from viewing the major role of evaluation as providing proof, to one of evaluation as assisting program developers to plan and deliver the best possible program interventions. Evaluations can be likened to problem solvers in this scenario. As indicated above, recognition that an evaluation agenda can change during a study means that evaluators must adopt more flexible procedures, in which questions are identified and answered, leading to further questions. This moves the evaluator more into the role of an investigator.

The investigatory process can be conceptualised as involving both exploratory and confirmatory phases. The evaluator needs to select from a range of methods, some of which collect qualitative data or a combination of qualitative and quantitative information.

It is important to note that while investigatory methods signify a shift in methodology, the principles which applied to data analysis mentioned in the previous section continue to apply. These are to ensure that the procedures used in the data collection and analysis phases are valid and reliable, and follow acceptable procedures. More will be said about methodology designed to meet these requirements below.

One interesting development in the move from evaluation as proof to evaluation as investigation is an **explicit** acknowledgement that the personal characteristics of the evaluator are important in data collection and analysis. That is, in order to carry out a sound evaluation, one needs sound mental powers as well as valid and reliable methodological tools. Smith (1992) suggests that an evaluator needs to bring four essential aspects to data collection and analysis in an investigative evaluation. These are:

- knowledge, about the program under study, the local context, and the broader social and political context in which it operates
- observation, knowing what to look for, the ability to receive and interpret, and to recognise the meaning and significance of what is seen
- reasoning, the ability to build and sustain a consistent line of argument about the program and its impact
- intuition, the use of insight and instinct about the program and its possible effects.

F. Information Analysis Techniques

As indicated earlier, data analysis involves making sense of the collected information in a way that enables the evaluator to develop conclusions or findings in which the stakeholders have confidence.

Summary statistics are a form of data reduction which makes numbers more manageable and helps to make sense of the data. Two ways in which statistical techniques can be useful include:

- to describe a situation. For example, a large set of scores on a test can be described by a variety of 'descriptive statistics'. The distribution of scores might be presented graphically, and then described by summary statistics which measure their range, their average (or mean) and whether they cluster closely around the average or whether they are spread out (i.e. the standard deviation from the mean).

- to answer key evaluation questions. For example, to investigate whether males and females respond differently to specific components of an adult literacy program. There are two questions that might serve as a starting point: Were the means (or averages) of the two groups different on a particular test ? and 'Were the standard deviations different?' Are these differences statistically significant ? Such information might lead to a set of hypotheses about why the differences might have occurred. Alternatively, it is possible to focus on patterns

> in the data by examining the relationships between variables. For example, it may be possible to establish a relationship between the failure to complete a literacy course and ethnic background. The statistic used to describe the strength of the relationship between two measures is called the correlation coefficient. It is often used in prediction, in that once a person's score on one of the measures is known, then it is possible to predict what his/her score might be on the other. (Fitzgibbon and Morris, 1987:78)

Further Reading:

Boruch, R., & Rindskopf, D. (1984) Data Analysis. In L. Rutman (ed.) Evaluation Research Methods (1977), Sage Publications, Newbury Park, CA.

de Vaus, D.A. (1990) Surveys in Social Research, Allen & Unwin Australia Pty. Ltd., North Sydney, NSW.

Fitzgibbon, C.T. & Morris, L.L. (1987) How To Analyze Data, Sage Publications, Newbury Park, CA.

Fowler, F.J. (1984) Survey Research Methods, Sage Publications, Newbury Park, CA.

Johnson, A.G. (1977) Social Statistics Without Tears, McGraw Hill, New York.

Loether, H.J. & McTavish, D.G. (1974) Descriptive Statistics for Sociologists, Allyn & Bacon, Boston.

While the rules for analysing data in numerical form have been reasonably well developed, there have been few systematic guidelines or shared rules for the analysis of qualitative data i.e. data in the form of words. Miles & Huberman (1984) suggest the following six major methods that might be useful for analysis **which takes place in conjunction with data collection:**

- Contact Summary Sheet. A contact summary is a single sheet with a series of focusing questions about a particular field contact. The questions provide a useful structure for developing a brief overall summary of important observations made in the field. Similarly, a document summary form may also be helpful in identifying key documents, explaining their significance and providing content summaries. Contact summary sheets include information about the key people, events or situations involved, the main issues which have emerged and their relevance to the focus of the study, and areas requiring further investigation and information collection (p.50).

- Codes and Coding. 'A code is an abbreviation or symbol applied to a segment of words - most often a sentence or paragraph of transcribed field notes - in order to classify the words. Codes are categories. They usually derive from research questions, hypotheses, key concepts, or important themes. They are retrieval and organizing devices that allow the analyst to spot quickly, pull out, then cluster all the segments relating to the particular question, hypothesis, concept, or theme. Clustering sets the stage for analysis.' (p.56)'. Coding is also discussed in 'Observation Forms', Chapter 7.

- Pattern Coding. After coding of segments of data has been completed, it is then possible to identify emerging patterns and group the summaries into overall themes. This process of coding permits a large amount of qualitative data to be reduced to a manageable level, it involves the researcher in analysis during data collection, and where there is more than one site and more than one fieldworker, it provides the basis for cross-site analysis (p.68).

- Memoing. 'Memos are always conceptual in intent. They do not just report data, but tie different pieces of data together in a cluster.' (p.69). Memos are reflective remarks about what is happening, and involve making connections between what is being observed in the field.

- Site Analysis Meeting. Where there is more than one site and more than one fieldworker in a study, it is important to have site-meetings to exchange information and ideas. Often the exchange is more effective if it is structured by a series of headings or key questions to which each fieldworker can respond (p.72).

- Interim Site Summary. This is an overall summary of the findings - a first attempt to provide a coherent account. It should highlight those areas or issues that need to be researched further and the kinds of data that still need to be collected (p.75).

Once data collection is substantially complete, various forms of display may be used to enable the evaluator to reduce the data and so lead to conclusion drawing. One form of data display which is extremely useful is a matrix display. The example in Figure 8.2 shows how data collected from site visits can be condensed into this format. In this case, the evaluation involved an investigation of innovatory use of new computer technology at five school sites in Victoria. Data were collected by interview, analysis of documentation and visits to the sites.

From an observation of the data, seven themes emerged which appear as row headings. The emergence of these themes depended (in Smith's terms above) on the information (knowledge) assembled by the evaluator and on his observation of patterns in the raw data. Entries in the matrix summarise the presence of these themes at each school, the names of which form the column headings. Note that, in investigatory enquiry, successive rounds of data collection are possible. In this example, the development of the matrix enabled the evaluator to contact the schools a second time to 'fill in' some of the matrix entries not tapped during the first round of data collection. The completed matrix became the basis of a comprehensive written and oral report to stakeholders (Owen, 1993b).

G. Good Practice In Information Collection

Refer to Chapter 10 for a discussion of good practice standards related to information collection.

Further Reading

de Vaus, D.A. (1990) Surveys in Social Research, Allen & Unwin Australia Pty. Ltd., North Sydney, NSW.

Miles, M.B. & Huberman, A.M. (1984) Qualitative Data Analysis A Sourcebook of New Methods, Sage Publications, Beverly Hills, California.

Fowler, F.J. (1984) Survey Research Methods, Sage Publications, Beverly Hills, California.

References

Challender, J. (1993) 'A DEET Case Study - The Advantages of Conducting a Pilot Study to Determine How Best to reach Employer Clients', Australasian Evaluation Conference, June.

de Vaus, D.A. (1990) Surveys in Social Research, Allen & Unwin Australia Pty. Ltd., North Sydney, NSW.

Fitzgibbon, C.T. & Morris, L.L. (1987) How To Analyze Data, Sage Publications, Newbury Park, London.

Fowler, F.J. (1984) Survey Research Methods, Sage Publications, Beverly Hills, California.

Henersen, M.E., Morris, L.L. & Fitz-Gibbon, C.T. (1987) How to Measure Attitudes, Newbury Park, California.

Kosecoff, J, & Fink, A. (1982) Evaluation Basics, A Practitioners Manual, Sage Publications, London and New Delhi.

Miles, M.B. & Huberman, A.M. (1984) Qualitative Data Analysis A Sourcebook of New Methods, Sage Publications, Beverly Hills, California.

Owen, J.M. (1993a) Program Evaluation Forms and Approaches, Allen & Unwin Pty. Ltd., St. Leonards, NSW.

Owen, J.M. (1993b) 'The Potential of Talking Books for Rural Schools in Victoria'. A report to the Country Education Project. Centre for Program Evaluation, The University of Melbourne.

Patton, M.Q. (1990), Qualitative Evaluation and Research Methods, 2nd Ed. Sage Publications, Newbury Park, London.

Smith, N.L. (1992) 'Aspects of Investigative Inquiry in Evaluation', New Directions in Program Evaluation, 56, 3-13, Winter, 1992.

FIGURE 8.1: PURPOSEFUL SAMPLING STRATEGIES

Type	*Purpose*
B. Purposeful sampling	Selects information-rich cases for in-depth study. Size and specific cases depend on study purpose.
1. extreme or deviant case sampling	Learning from highly unusual manifestations of the phenomenon of interest, such as outstanding successes/notable failures, top of the class/dropouts, exotic events, crises.
2. intensity sampling	Information-rich cases that manifest the phenomenon intensely, but not extremely, such as good students/poor students, above average/below average.
3. maximum variation sampling - purposefully picking a wide range of variation on dimensions of interest	Documents unique or diverse variations that have emerged in adapting to different conditions. Identifies important common patterns that cut across variations.
4. homogeneous sampling	Focuses, reduces variation, simplifies analysis, facilitates group interviewing.
5. typical case sampling	Illustrates or highlights what is typical, normal, average.
6. stratified purposeful sampling	Illustrates characteristics of particular subgroups of interest; facilitates comparisons.
7. critical case sampling	Permits *logical* generalisation and maximum application of information to other cases because if it's true of this one case it's likely to be true of all other cases.
8. snowball or chain sampling	Identifies cases of interest from people who know people who know people who know what cases are information-rich, that is, good examples for study, good interview subjects.
9. criterion sampling	Picking all cases that meet some criterion, such as all children abused in a treatment facility. Quality assurance.
10. theory-based or operational construct sampling	Finding manifestations of a theoretical construct of interest so as to elaborate and examine the construct.
11. confirming and disconfirming cases	Elaborating and deepening initial analysis, seeking exceptions, testing variation.
12. opportunistic sampling	Following new leads during fieldwork, taking advantage of the unexpected, flexibility.
13. random purposeful sampling (still small sample size)	Adds credibility to sample when potential purposeful sample is larger than one can handle. Reduces judgment within a purposeful category. (Not for generalisations or representativeness.)
14. sampling politically important cases	Attracts attention to the study (or avoids attracting undesired attention by purposefully eliminating from the sample politically sensitive cases).
15. convenience sampling	Saves time, money, and effort. Poorest rationale; lowest credibility. Yields information-poor cases.
16. combination or mixed purposeful sampling	Triangulation, flexibility, meets multiple interests and needs.

Source: Patton, M.Q. (1990) Qualitative Evaluation and Research Methods, Sage Publications, Newbury Park, London, p.182-183

FIGURE 8.2: MATRIX DISPLAY

Attachment A	Euroa	Sea Lake	Boort	Minyip	Whitfield
Curriculum content	music, year 12 oral comp. thematic units	LOTE, German, English, Japanese-English	Mathematics	Maths, language, geography, music, LOTE	LOTE, Italian
teaching styles	independent work for year 12 students • student planned unit	class use with small groups and in out of class time	reinforcement of content taught in whole class group	various styles, independent work for advanced students	• use of contextual class materials • sharing of student work across sites in the cluster
student activities	• independent study • student developed units • whole class	cooperative production of materials		• self assessment • sound dictionary use • independent extension work	• student product development (book)
student assessment	• year 12 cat in music	word and phrase construction check	sheets for testing mathematics achievement	sheets for testing achievement, more complex tests of higher skills	
affective outcomes	increased enrolments and interest in music	excitement among year 6-8 students		novelty value to students	increased interest in Italian across cluster sites
links to other staff					encouragement and support for other staff to teach LOTE
publicity		• importance of LOTE in the curriculum • interest among school community			

Source: Owen, J.M. (1993) The Potential of Talking Books for Rural Schools in Victoria. A report to the Country Education Project. Centre for Program Evaluation, The University of Melbourne.

Chapter 9 REPORTING IN EVALUATION

It is essential that evaluations impact on decision making. In this chapter we present guidelines to ensure that this happens. Evaluative information could relate to a range of aspects of a given program. Evaluation reporting may include information about the progress of students but this is not necessarily the case. As we shall see below, a major factor in determining what information presented is to determine the needs of the evaluation audience.

A. Determining Audience Needs

As indicated in *overview framework* Figure2.4 (in Chapter 2), the first step in planning evaluation reporting is to identify the major audiences and their information needs. Reporting designs should take into account:

- whether the primary role of the evaluation is to provide feedback, aid decision-making, initiate improvement or some combination of all three
- the preferences of the primary audience about the format and language used in presenting findings
- the time-frame for interim and final decisions about the program under review
- interim and final reporting deadlines imposed as part of the evaluation design.

There are a number of reasons why responding to audience needs may become a complex task:

- some audiences do not know what they need. It is frequently necessary for those conducting the evaluation to help program staff clarify their objectives and define their information needs.
- different audiences often require different information. Consider for example, an evaluation of the effectiveness of a staff training program. Whereas program staff and clients might find interview responses and anecdotal reports acceptable as evidence of program effects, external funding bodies might require test data that meets certain standards of validity and reliability (Morris, et.al. 1987:14)
- information needs can change during the course of an evaluation (Morris, et.al. 1987:15). This can occur particularly when a new and developing program is being evaluated for the purpose of improvement. Issues which were not foreseen at the outset of the study emerge over time and warrant attention.

Figure 2.4 summarises examples of possible communication forms and their relevance for a range of audiences.

B. Reporting Options

It is important to realise that final reporting of an evaluation need not necessarily take the form of a written document. With increasing concern about the lack of impact of evaluation findings, the emphasis has moved away from long formal written reports to strategies which allow clients and audiences to interact with those conducting the evaluation. It is the responsibility of the evaluator to ensure that audiences comprehend the key findings of an evaluation. This is an integral part of an evaluator's work. Ideally, the ways in which the evaluator disseminates the findings and conclusions should be negotiated during the planning phase of the evaluation. Some of the reporting options which need to be considered during that phase include:

a. written versus oral reports (with or without visual aids)

- written reports. Examples are:
    - a document including a summary of the key findings with the detailed findings in supporting appendices
    - a number of individual summaries of key findings directed towards specific audiences
    - a press release
    - an informal letter.
- oral reports. Examples are:
    - press conferences
    - informal discussions
    - verbal presentations to select groups who might become actively involved in the analysis and interpretation of the findings
    - reporting sessions based around a series of graphics, charts or displays
    - a short video- or audio-tape designed to convey the findings in the absence of the evaluator
    - a debate or advocacy-adversary analysis of the findings allowing alternative opinions to be expressed in a group forum (Patton, 1986:272-273).

Some of the strategies identified above involve stakeholders `in the processes of interpreting data, making judgments, drawing conclusions, and generating recommendations'. Patton (1986:272) sees this as an essential part of the educational role integral to the evaluation process. Stakeholder involvement in the course of the evaluation is a means of gaining support for the findings and eliminating any surprises at the end of the evaluation process.

b. progress versus final reports

Both the content and the timing of information needs may change in the course of an evaluation. Under these circumstances, it is not uncommon for decisionmakers to make requests for short interim reports to be provided.

Frequently, interim reports will be part of the reporting design agreed to at the outset of evaluation. The use of such reports has a number of advantages:

- a series of short reports spaced out over time are more likely to be read carefully by most audiences than one long final report at the end of the study. Decisionmakers in senior positions, in particular, lack the time to read large detailed reports.
- a short report can focus on specific issues/areas within the evaluation and hence there is a greater chance of an essential message being conveyed.
- short reports can be used to disseminate various discrete findings about a particular stage of the evaluation, findings required by decisionmakers at a particular time.
- the workload of those conducting the evaluation is spread out over a period of time.

c. substantive (main report) versus secondary (such as technical details of data management)

Depending on the audience, a main report without the technical details regarding data management may be appropriate. For example, community groups are likely to be most interested in the findings expressed in simple direct reports and are not likely to want the technical details. Alternatively, service providers involved in implementing the program under scrutiny may appreciate a full description of how the program evaluation was conducted, methods of data management, etc.

d. summary versus main report

Frequently, a main report is accompanied by a summary of key findings, but there are times when this summary will stand alone. It may be accompanied by graphic displays or structured by a series of focussing questions.

It is not uncommon for the final report of an extensive evaluation study to be a large and comprehensive document. While decisionmakers may require the complete report, a summary of key findings is useful for disseminating information to a range of wider audiences. For example, while a regional co-ordinator might require a detailed main report, it would be costly and probably ineffective to reproduce many copies of the report as a means of informing service providers of the findings. Distributing copies of the summary of key findings would not only be less costly but is likely to be a more successful means of communication with those out in the field.

e. formal versus informal

While reports expressed in formal language are often required, particularly by external funding bodies, there is some evidence to suggest that less formal reports are more easily understood and likely to increase the impact of the findings (Owen, 1993:81). The content of formal evaluation reports vary according to the evaluation and its context. Nevertheless, as a starting point in writing a formal report, the following main sections may prove useful (Morris, et.al., 1987:78-89):

- A Front Cover, which gives the title of the program, its location, the name of the evaluator(s), the period covered by the report and the date it was submitted.
- A Summary of Key Findings/Recommendations
- Background Information About the Program including information about the origins of the program, its objectives, a description of program clientele, characteristics of program materials, activities and administrative arrangements, and details of staffing and others involved in program implementation.
- A Description of the Evaluation Study, including key evaluation purposes, evaluation design, outcome and implementation measures.
- Evaluation Findings, including, for example, summaries of questionnaire data, anecdotal evidence, testimonials, excerpts from interviews, or statistical data presented in graphs or tables.
- A Discussion of the Findings where the implications of the findings as given in the previous section are discussed, and the relevance of the findings to the key purposes of evaluation are highlighted.
- The Costs and Benefits of the Program (Optional). This section is not always included in evaluation reports. If it is, however, it usually involves a justification of the approach taken to the cost/benefit analysis and a discussion attempting to weigh up program costs against perceived benefits.

- Conclusions and Recommendations. Not all evaluation reports include recommendations. If they do, they need to be expressed and presented with great care (see below for further discussion on making recommendations).

f. descriptive versus recommendatory

It is important to make the distinction between four different kinds of information in evaluation reports:

- findings which refer to the data arising from the evaluation, the empirical results.

- interpretations or conclusions which are explanations offered about the findings. These may involve speculation about the meaning of the data, reasons for the findings, and interrelationships observed.

- judgements in which values are brought to bear on the data; `specific criteria applied to the findings stating that they are "good" or "bad"; "positive" or "negative"; "in the desired direction" or "in an undesired direction"; "above expectations", "in line with expectations," or "below expectations."'

- recommendations which include `suggested courses of action; proposed changes in the program or things to be maintained as they are in the program; and advice to funders, program administrators, program staff, and others about how to improve the program based on findings, interpretations and judgments.' (Patton, 1982:273)

As indicated earlier, the information contained in an evaluation report should be negotiated early in the evaluation process. At the simplest level, decisionmakers may only require a description of the findings. At the next level of complexity, they may want an analysis of the findings but without any interpretation placed on that analysis. In some instances, reporting will include interpretations or judgements (or both) based on the findings. At the highest level of complexity, reports will include description, analysis, interpretations/judgements and a series of recommendations (Patton, 1982:283-284).

Ways in which recommendations are presented in final reporting will have a significant bearing on the impact they have on their audience. For example, close attention to the order in which recommendations are presented is important. Recommendations central to the evaluation purpose can be lost if included amongst a long undifferentiated list of recommendations at the end of a report. There are a number of ways of organising recommendations depending on the purpose and context of the evaluation and the evaluation audience. Below are some suggested alternatives for grouping `like recommendations' together:

- recommendations of critical importance to the evaluation as distinct from recommendations of lesser importance, OR

- recommendations strongly supported by the data as distinct from recommendations less well supported or subject to some disagreement, OR

- recommendations which relate to different aspects of program operations. For example:
    - management and day-to-day issues
    - basic program policies
    - program funding
    - program content

494

495

  - program processes
  - program organisation, OR

- recommendations which relate to different stakeholder groups. For example:
  - funding agencies
  - program administrators
  - program staff
  - program participants, OR

- recommendations involving different implementation time-frames. For example, those that can be implemented:
  - immediately
  - in six months time
  - in twelve months time (Patton, 1982:278-279).

Recommendations can be presented as a series of options. Decision-makers often value a full discussion of each option or course of action that follows from the findings and may be more receptive to pursue the evaluator's preferred option if they have had the chance to consider the alternatives.

Care must be taken to avoid making recommendations which relate to factors outside the control of the decisionmakers. When making recommendations which involve politically sensitive issues or which call for radical change the response of the various evaluation audiences needs to be anticipated and strategies prepared to cater for their response.

Recommendations can be drawn from a variety of sources: earlier studies of the same program, successful sites, similar programs, the ideas of program staff, speculation by decision-makers. Where possible, it is often helpful to identify the future implications of each recommendation, outlining an implementation strategy and the resources required (Hendricks & Papagiannis, 1990:121-125).

`Writing recommendations is a strategic and creative process. It is the culmination of the evaluation effort. The recommendations represent the most visible linkage between the data and subsequent utilization of the evaluation findings' (Patton, 1982:289).

C. Good Practice In Evaluation Reporting

Refer to Chapter 10 for a discussion of good practice standards related to evaluation reporting.

References

Hendricks, M. & Papagiannis, M. (1990) `Do's and Don'ts for Offering Effective Recommendations', Evaluation Practice, Vol. 11, No. 2, pp.121-125.

Morris, L.L., Fitzgibbon, C.T., & Freeman, M.E. (1987) How to Communicate Evaluation Findings, Sage Publications, Newbury Park, California.

Owen, J.M. (1993) Program Evaluation Forms and Approaches, Allen & Unwin Pty. Ltd., St. Leonards, NSW.

Patton, M.Q. (1982) Practical Evaluation, Sage Publications, Newbury Park, California.

Patton, M.Q. (1986) Utilization-Focused Evaluation (2nd Ed.) Sage Publications, Newbury Park, California.

Further Reading:

Morris, L.L., Fitzgibbon, & Freeman, M.E. (1987) How to Communicate Evaluation Findings, Sage Publications, Newbury Park, California.

Patton, M.Q. (1988) The future and evaluation, Evaluation Practice, 9 (4), 90-93.

Roberts-Gray, C., Buller, A., & Sparkman, A. (1987) Linking data with action: Procedures for developing recommendations. Evaluation Review, 11, 678-684.

Chapter 10 GOOD PRACTICE IN EVALUATION

Evaluation practice must make an effective contribution to decision making about social and educational programs. To make such a contribution evaluations must be done within an acceptable ethical framework. During the development of this project, some of the assumptions underlying acceptable evaluation practice have been developed.

These include that ALBE evaluations must:

- be 'owned' by the field
- identify legitimate stakeholders (including the Commonwealth) and include them in decision making about the evaluation
- make evaluation processes simple and accessible to stakeholders
- focus on critical areas of adult basic education, and
- be cost effective.

These issues are discussed and expanded upon in this chapter.

Stakeholder Involvement in the Evaluation Process

As indicated previously, involving stakeholders in an evaluation significantly increases the likelihood that the findings will have an impact on the program. Stakeholder involvement helps to ensure that the issues selected and the questions that focus the evaluation are relevant and important, and hence that the findings are useful. Good or 'utilization-focused' evaluation questions directly reflect the information needs of stakeholders. The primary users of the information should want and need the information for their own purposes, not just for someone else. They should care about the answer to the question and be able to envisage how the information will lead to further action (Patton, 1978:69-70).

Many evaluations are commissioned by groups, and it is an important skill of an evaluator to determine the group agenda and to encourage the group commitment to use of the findings. Patton and his associates (1978:64) found that ultimately the single most important factor in determining the use of evaluations was the 'personal factor' i.e. 'the presence of an identifiable individual or group who personally cared about the evaluation and the information it generated'.

Primary 'use' of an evaluation may be thought of as comprehending the findings. However, the process itself can be seen as a learning experience for participants providing them with the opportunity to:

- reflect on the work that they are doing
- conceptualise issues or aspects of a program in new ways
- identify gaps in their existing knowledge about an issue or a program
- understand more about the complexity of program development and human service delivery
- acquire knowledge and skills in evaluation (Patton, 1982:298-303).

Identifying Evaluation Priorities

Within complex organisations delivering ALBE, or programs offered within ALBE, decisions must be made about what should be evaluated given the limited resources available.

The following steps provide a structure for identifying evaluation priorities:

- **make a list of potential issues or programs which warrant evaluation.** Suggestions might come from a variety of sources: program staff or clients, reviews of program plans, managers at regional and state level, etc.

- **eliminate any nominated areas of evaluation that are non-issues or could be handled in some other way.** This may require some preliminary data gathering about the issues from a range of key informants.

- **assess the feasibility of undertaking evaluation in the remaining areas listed,** with a view to eliminating those which are too difficult to evaluate given existing conditions and resources. The following questions can be used to assess degree of difficulty:
  - how much time would be required to complete the evaluation?
  - what would be the costs, including staffing?
  - what level of resistance is likely to be encountered, for example, from program staff ?
  - how well-defined is the program? Are its goals and performance criteria clear?
  - what is the program's stage of development? If an impact evaluation is contemplated, is it possible for such an evaluation to be undertaken now?
  - is it possible to achieve valid and reliable measurements of those aspects of the program which are to be studied?
  - what political, administrative or legal constraints might affect the collection of data or the ability to implement recommendations?

- **establish priorities within the remaining areas to be evaluated,** using the following questions:
  - what is the nature and level of concern about the issue or program proposed for evaluation?
  - what is the degree of urgency associated with the issue or program proposed for evaluation?
  - what is the level of impact of the issue on the program, or the level of impact of the program on the organisation as a whole?
  - what are the consequences to the program or the organisation as a whole of not carrying out an evaluation?
  - what is the significance of the issue or program in terms of financial investment?
  - how quickly is the evaluation information required?
  - how difficult is it to carry out the evaluation? (see questions given above).

Muscatello (1988) devised a series of rating scales which allow the degree of difficulty of a proposed evaluation area to be quantitatively assessed. This summary assessment rating is then considered alongside the remaining criteria in developing final evaluation priorities. This systematic approach to determining priorities is designed for internal evaluation units operating within large complex organisations (see references at end of this chapter).

Internal Evaluations: A Set of Guidelines

Internal evaluations are undertaken by 'insiders' i.e. by those who belong to the organisation in which the evaluation is occurring.

Internal evaluation is more effective when:

- program staff are involved as much as possible in the planning and implementation of the study.

- an individual (or two individuals operating as a team) are given clearly defined responsibilities for the management of the evaluation.

- there is agreement (as far as is possible) on the evaluation plan.

- the evaluation team is given access to resources to support them in the collection and analysis of information. This may involve external technical advice on data management and analysis.

- interim reports on the progress of the evaluation are provided.
- the findings are used to make decisions about program changes.
- action plans are developed for implementing program changes. These plans should identify the individuals responsible for seeing that action is taken within designated time-frames.

'Criticisms of internal evaluations include their lower credibility as seen from outside the organisation, the lack of expertise in data management, and the tendency for staff to become exhausted by the additional effort needed to undertake the task. These problems can be ameliorated by containing the study and by providing resources for its management. In addition, it is increasingly clear that organisations should hire staff with appropriate, systematic training to assist with evaluation planning and execution.' (Owen, 1993:39-40)

Standards for Evaluators

In 1981, the Joint Committee on Standards for Educational Evaluation (USA), composed of representatives from twelve professional societies with interests in evaluation, published Standards for Evaluation of Educational Programs, Projects and Materials (Stufflebeam, 1986). The 30 standards are grouped in the following four areas in order of priority:
- utility
- feasibility
- propriety
- accuracy.

The committee's position with respect to the priority order of these standards is explained by Dan Stufflebeam, Chair of the Evaluation Standards Committee:

> '... an evaluation should not be done at all if there is no prospect for its being useful to some audience. Second, it should not be done if it is not feasible to conduct it in political terms, or practicality terms, or cost-effectiveness terms. Third, ... it should (not) be done if we cannot demonstrate that it will be conducted fairly and ethically. Finally, if we can demonstrate that an evaluation will have utility, will be feasible and will be proper in its conduct then they said we could turn to the difficult matters of the technical adequacy of the evaluation.' (Stufflebeam, 1980:90)

Utility Standards

The utility standards are intended to ensure that an evaluation will serve the practical information needs of given audiences. These standards are:

U1 Audience Identification
Audiences involved in or affected by the evaluation should be identified so that their needs can be addressed.

U2 Evaluator Credibility
The persons conducting the evaluation should be both trust-worthy and competent to perform the evaluation, so that their findings achieve maximum credibility and acceptance.

U3 Information and Scope
Information collected should be of such scope and selected in such ways as to address pertinent questions about the object of the evaluation and should be responsive to the needs and interests of specified audiences.

U4 Valuation Interpretation
The perspectives, procedures and rationale used to interpret the findings should be carefully described, so that the bases for value judgements are clear.

U5 Report Clarity

The evaluation report should describe the object being evaluated and its context, and the purposes, procedures and findings of the evaluation, so that the audience will readily understand what was done, why it was done, what information was obtained, what conclusions were drawn, and what recommendations were made.

U6 Report Dissemination
Evaluation findings should be disseminated to clients and other right-to-know audiences, so that they can assess and use the findings.

U7 Report Timelines
Release of reports should be timely, so that audiences can best use the reported information.

U8 Evaluation Impact
Evaluations should be planned and conducted in ways that encourage follow through by members of the audiences.

Feasibility Standards

The feasibility standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic and frugal. These standards are:

F1 Practical Procedures
The evaluation procedures should be practical, so that disruption is kept to a minimum and needed information can be obtained.

F2 Political Viability
The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation can be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or apply the results can be averted or counteracted.

F3 Cost Effectiveness
The evaluation should produce information of sufficient value to justify the resources expended.

Propriety Standards

The propriety standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results. These standards are:

P1 Formal Obligation
Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that the parties are obliged to adhere to all the conditions of the agreement - or formally to renegotiate it.

P2 Conflict of Interest
Conflict of interest, frequently unavoidable, should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results.

P3 Full and Frank Disclosure
Oral and written evaluation reports should be open, direct and honest in their disclosure of pertinent findings, including the limitations of the evaluation.

P4 Public's Right to Know
The formal parties to an evaluation should respect and assure the public's right to know within the limits of other related principles and statutes, such as those dealing with public safety and the right to privacy.

P5 Rights of Human Subjects
Evaluations should be designed and conducted so that the rights and welfare of the human subjects are respected and protected.

P6 Human Interactions

Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation.

P7 Balanced Reporting

The evaluation should be complete and fair in its presentation of strengths and weaknesses of the object under investigation, so that strengths can be built on and weaknesses addressed.

P8 Fiscal Reporting

The evaluator's allocation and expenditure of resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible.

Accuracy Standards

The accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features of the object being studied that determine its worth or merit. These standards are:

A1 Object Identification

The object of the evaluation (program, project, material) should be sufficiently examined, so that the form(s) of the object being considered in the evaluation can be clearly identified.

A2 Context Analysis

The context in which the program, project or material exists should be examined in detail, so that its likely influence on the object can be identified.

A3 Described Purposes and Procedures

The purposes and procedures of the evaluation should be monitored and described in enough detail to be identified and assessed.

A4 Defensible Information Sources

The sources of information should be described in enough detail to allow the adequacy of the information to be assessed.

A5 Valid Measurement

The information-gathering instruments and procedures should be chosen or developed and then implemented in ways that will assure that the interpretation arrived at is valid for the given use.

A6 Reliable Measurement

The information-gathering instruments and procedures should be chosen or developed and then implemented in ways that will assure that the interpretation arrived at is reliable for the intended use.

A7 Systematic Data Control

The data collected, processed and reported in an evaluation should be reviewed and where necessary corrected, so that the results of the evaluation will not be flawed.

A8 Analysis of Quantitative Information

Quantitative information in an evaluation should be appropriately and systematically analysed to ensure supportable interpretation.

A9 Analysis of Qualitative Information

Qualitative information in an evaluation should be appropriately and systematically analysed to ensure supportable interpretation.

A10 Justified Conclusions

The conclusions reached in an evaluation should be explicitly justified so that audiences can assess them.

A11 Objective Reporting

The evaluation procedures should provide safeguards to protect the findings against distortion by the personal feelings and biases of any party to the evaluation.

> **'Taken together, the standards express a commitment to fulfilling the promise of evaluation, namely, providing high-quality information for use in improving and making decisions about programs.' (Patton, 1982:299)**

References

Muscatello, D.B. (1988) 'Developing an Agenda That Works: The Right Choice at the Right Time', in J.A. McLaughlin, L.J. Weber, R.W. Coverts, and R.B. Ingle (eds) Evaluation Utilization, New Directions for Program Evaluation, No.39, San Francisco: Jossey-Bass, Fall, 1988.

Owen, J.M. (1993) Program Evaluation Forms and Approaches, Allen & Unwin Pty. Ltd., St. Leonards, NSW.

Patton, M.Q. (1978) Utilization-Focused Evaluation, Ist ed. Sage Publications Inc., Beverly Hills, California.

Patton, M.Q. (1982) Practical Evaluation, Sage Publications, Newbury Park, California.

Patton, M.Q. (1986) Utilization-Focused Evaluation, 2nd Ed. Sage Publications, Newbury Park, California.

Stufflebeam, D.L. (1980) An Interview with Daniel Stufflebeam, Educational evaluation and policy analysis, (1980), 2(4).

Stufflebeam, D.L. (1986) 'Standards of Practice for Evaluators', Paper given at the Annual Meeting of the American Evaluation Research Association, San Francisco, April.

CE 070106

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE
(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Professional Development for Program Evaluation: Evaluation for Adult Literacy and Basic Education Programs	
Author(s): Rex Stoessiger and Ray Pugh	
Corporate Source:	Publication Date: 1995

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

☑ ← **Sample sticker to be affixed to document**

Check here
Permitting microfiche (4" x 6" film), paper copy, electronic, and optical media reproduction.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

Level 1

Sample sticker to be affixed to document → ☐

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

Level 2

or here
Permitting reproduction in other than paper copy.

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature:	Position: ARIS OFFICER
Printed Name: Jan Kindler	Organization: ARIS/NLLIA
Address: Level 9 300 Flinders street Melbourne VIC 3001	Telephone Number: (03) 9614 0255
	Date:

OVER

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information reguarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

[redacted]/Distributor: ARIS/NLLIA

Address: Level 9
300 Flinders Street
Melbourne VIC 3001

Price Per Copy: $40 00

Quantity Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name and address of current copyright/reproduction rights holder:

Name:

Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

Associate Director for Database Development
ERIC Clearinghouse on Adult, Career, and Vocational Education
Center on Education and Training for Employment
1900 Kenny Road
Columbus, OH 43210-1090

If you are making an unsolicited contribution to ERIC, you may return this form (and the document being contributed) to:

(Rev. 9/91)